PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/82, 15/29, C07K 14/415, C12N 5/10, A01H 1/00 — A1

(11) International Publication Number: **WO 00/32799**

(43) International Publication Date: 8 June 2000 (08.06.00)

(21) International Application Number: PCT/US99/28103

(22) International Filing Date: 24 November 1999 (24.11.99)

(30) Priority Data:
60/109,892 25 November 1998 (25.11.98) US

(71) Applicant: CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).

(72) Inventor: CHAUDHURI, Sumita; 1125 Juniper Place, Davis, CA 95616 (US).

(74) Agent: SCHWEDLER, Carl, J.; Patent Department Central, Monsanto/G.D. Searle, P.O. Box 5110, Chicago, IL 60680-5110 (US).

(81) Designated States: CA, JP, MX, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS FOR TRANSFORMING PLASTIDS

(57) Abstract

An improved method is provided for the transformation of a plant cell plastid. The improved method allows for the increased efficiency of the foreign DNA penetrating the plastid membrane. The method generally involves the use of a plant tissue source having an altered plastid morphology in plastid transformation methods. The present invention finds use in plastid transformation methods for a wide variety of plant species.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

METHODS FOR TRANSFORMING PLASTIDS

FIELD OF THE INVENTION

This invention relates to the application of genetic engineering techniques to plants. More specifically, the invention relates to methods for the transformation of plant cell plastids.

BACKGROUND

The plastids of higher plants are an attractive target for genetic engineering. Plant plastids (chloroplasts, amyloplasts, elaioplasts, chromoplasts, etc.) are the major biosynthetic centers that in addition to photosynthesis are responsible for production of industrially important compounds such as amino acids, complex carbohydrates, fatty acids, and pigments. Plastids are derived from a common precursor known as a proplastid and thus the plastids present in a given plant species all have the same genetic content. Plant cells contain 500-10,000 copies of a small 120-160 kilobase circular genome, each molecule of which has a large (approximately 25kb) inverted repeat. Thus, it is possible to engineer plant cells to contain up to 20,000 copies of a particular gene of interest which potentially can result in very high levels of foreign gene expression.

Current plastid transformation methods are inefficient, as such there is need for constructs and methods which improve plastid transformation.

SUMMARY OF THE INVENTION

By this invention, methods which allow for the improved transformation of a foreign DNA into plant cell plastids are provided. Such methods generally involve

utilizing a plant tissue source which contains cells with an altered plastid morphology in the transformation methods. The alteration in the plant plastid morphology includes, *inter alia*, plastid size and number. By utilizing tissue derived from such plants in plastid transformation methods, efficiency of transformation of a foreign DNA into the plant cell plastid may be increased.

As exemplified herein, constructs useful for genetic engineering of plant cells to provide for a method of increasing plastid transformation efficiency are provided. The constructs include nucleic acid sequences coding for protein sequences involved in controlling division of plant cell organelles. The expression of such nucleic acid sequences in a plant cell provides for an altered number and/or size of the chloroplasts within the host cell.

DNA sequences, also referred to herein as polynucleotides, for use in transformation contain an expression construct comprising a promoter region which is functional in a plastid, and a DNA sequence encoding a gene involved in controlling the division of plant cell organelles.

Methods for the use of transformed plants with altered plastid morphology are described. Such methods include plant breeding or transformation methods to provide plant cells having both the nuclear and plastid constructs.

The present invention also provides methods for increasing the efficiency of chloroplast transformation. The method generally comprises transforming the plastids of a plant tissue which has been modified to have an altered number and/or size of plastids contained within the plant cell.

The present invention also provides a mechanism for enhancing the efficiency of chloroplast transformation in plant species.

The present invention also provides methods for improving the selectability of plant comprising, transforming a plant cell source having an altered plastid morphology with a construct comprising a promoter functional in a plant cell plastid operably associated with a nucleic acid sequence encoding a selectable marker. Selectable markers of interest in the present invention include herbicide tolerance genes such as glyphosate tolerance genes, and antibiotic resistance genes. Glyphosate tolerance genes include the CP4 gene from Agrobacterium.

Another aspect of the present invention provides methods for preparing a plant cell source with increased plastid transformation efficiency comprising, transforming a plant cell with a construct comprising a promoter functional in plant cell operably associated with a nucleic acid sequence encoding a FtsZ protein.

Also considered part of this invention are the plants and plant cells obtained using the methods described herein.

DESCRIPTION OF THE FIGURES

Figure 1 provides an amino acid sequence alignment of the Arabidopsis FtsZ1 (SEQ ID NO:2), the Brassica FtsZ1 (SEQ ID NO:6), the tobacco FtsZ1 (SEQ ID NO:9), the Soybean FtsZ1 (SEQ ID NO:72) and the corn FtsZ1 (SEQ ID NO:73) protein sequences.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the subject invention, methods are provided which allow for the improved transformation of a foreign DNA into plant cell plastids. Such methods generally involve utilizing a plant cell source which contains an altered plant plastid morphology. By utilizing tissue derived from such plants in plastid transformation methods, efficiency of transformation of a foreign DNA into the plant cell plastid can be increased.

In one embodiment of the instant invention, plant tissue containing altered plant plastid morphology is used for plastid transformation methods. Such alterations in plant plastid morphology include, but are not limited to, alterations in the plastid size, shape and number in respect to a wild-type plastid morphology from the target plant cell. In general, a wild-type plastid morphology consists of small, round organelles contained within the plant cell, depending on the species. Furthermore, a plant cell typically contains between about 50 and about 100 plastids.

The plant tissue source used in plastid transformation methods of the present invention contains an increase in the size of the plastids contained in the plant cells.

Such increases in the size of the plastids provides for a larger surface area for the foreign DNA to penetrate the plastid membrane during transformation.

The large plastids preferably contain approximately the same number of plastid genomes as would be contained in corresponding number of wild-type plastids. For example, in a wild-type plant cell containing 100 plastids per cell and 100 copies of the plastid genome in each plastid (a total of 10,000 copies of the plastid genome per cell), the corresponding mutant tissue source would preferably contain about the same number of plastid genomes, only contained in one, or several large plastid(s).

Alternatively, a plant tissue source with an increased number of plastids, with a corresponding reduced size, can also find use in the plastid transformation methods of the present invention.

As is understood in the art, additional methods for obtaining plants with alterations in the plastid size and number are known. The skilled artisan will recognize that a number of methods are available for providing for an alteration in plastid cell division. Such methods are described, for example, by Strepp, *et al.* (1998) *Proc. Natl. Acad. Sci. USA*, 95:4368-4373.

Cell division, also referred to as cytokinesis, has been the focus of studies in many organisms such as bacterial, fungal, and animal cells. Division of bacterial cells occurs through the formation of an FtsZ ring (also referred to as a Z ring) at the site of division (Lutkenhaus, *et al.* (1997) *Ann. Rev. Biochem.* 66:93-116). The positioning and formation of the Z ring acts to further drive septation (cytokinesis). The ring is composed of a tubulin-like FtsZ protein which has GTPase activity. Mutations in the *fts*Z gene in *E. coli* leads to the production of a temperature-sensitive filaments with regularly spaced nucleoids at certain temperatures (Lutkenhaus (1992) In *Prokaryotic Structure and Function: A New Perspective*, ed. S Mohan, C Dow, pp 123-152. Cambridge: Cambridge Univ. Press). Such mutations in bacteria leads to the inability to divide correctly.

The plant cell plastid as well as the mitochondria are derived from prokaryotic ancestors, and thus, the division apparatus of these organelles resembles that of bacteria. Recently, identification of *fts*Z related sequences in *Arabidopsis* and *Physcomitrella patens* have been reported (Osteryoung, *et al.* (1995) *Nature*, 376:473-

474; and Strepp, *et al.* (1998), *supra*). The protein encoded by the *Arabidopsis ftsZ* gene was found to be imported into the chloroplast and was therefore speculated to be a component of the plastid division machinery (Osteryoung, *et al.* (1995), *supra*). More recently, the involvement of FtsZ in plastid division was directly demonstrated. The disruption of the *ftsZ* gene in a lower plant, *Physcomitrella patens*, impeded plastid division, thereby giving rise to mutant cell lines with one or a few large plastids (Strepp, *et al.* (1998), *supra*).

The use of plants with an altered number and/or size of plastids containing one or few large plastids could therefore be used as targets for plastid transformation of any plant species. Such plants containing an altered size and/or number can be obtained using various methods, including mutagenesis, antisense suppression, or co-suppression. Methods for the mutagenesis of plant genomes are well known in the art, and include chemical, such as ethylmethane sulfonate (EMS) and nitrosoguanidine (NTG), as well as physical mutagenesis methods such as fast neutron bombardment.

Other means for obtaining a plant source with an alteration in the size and/or number of plastids contained in the cell are also contemplated. For example, tissue for use in the transformation methods of the present invention can be obtained from plants grown in culture conditions which provide for such altered plastid content. For example, tissue obtained from plants grown *in vitro* under culture conditions in which inhibitors of bacterial cell division, such as 5,5'-bis-(8-anilino-1-naphtalenesulfonate) (Yu, *et al.* (1998) *J. Biol Chem.* 273:10216-10222), are present, can be utilized as a cell source for the plastid transformation methods of the present invention.

In a preferred embodiment, such plants containing cells with an alteration in the size and/or number of plastids are generated by anti-sense expression of the FtsZ gene. Once plastid transformation is achieved and homoplasmic plants are identified, the anti-sense transgene can be eliminated by out-crossing and the wild-type condition of 50 to 100 plastids per cell restored. Similarly, plants regenerated from plastid transformed tissue containing an altered number and/or size of plastids from mutations can also be reverted to the wild-type plastid conditions using such out-crossing methods.

In the case of the use of culture conditions for obtaining plant cells with an altered number and/or size of plastids, wild-type plastids can be obtained by releasing the tissue from such culture conditions.

In another embodiment of the present invention, novel nucleic acid sequences are provided which encode proteins related to proteins involved in bacterial cell and plastid division.

In particular, novel nucleic acid sequences from *Arabidopsis*, soybean, corn, Brassica are provided which encode FtsZ related proteins. Such nucleic acid sequences find use in the preparation of DNA constructs. Such constructs find use in the production of plants with an altered number and/or size of chloroplasts.

The skilled artisan will recognize that other DNA sequences useful for the production of plants with an altered number and/or size of chloroplasts are available in the art. The sequences include but are not limited to, *fts*A, *fts*L, *fts*I, *fts*Q, *fts*N, *fts*W, *fts*K (Lutkenhaus, *et al.* (1997) *supra*), and the *arc* genes (Pyke, *et al.* (1992) *Plant Physiol.* 99:1005-1008; Pyke *et al.* (1994) *Plant Physiol.* 104:201-207; and Pyke (1997) *Am. J. Botany* 84:1017-1027).

In order to obtain additional *fts*Z sequences, a genomic or other appropriate library prepared from the candidate plant source of interest can be probed with conserved sequences from one or more plant and/or bacterial *fts*Z sequence(s) to identify homologously related sequences. Positive clones can be analyzed by restriction enzyme digestion and/or sequencing. When a genomic library is used, one or more sequences can be identified providing both the coding region, as well as the transcriptional regulatory elements of the *fts*Z gene from such plant source. Probes can also be considerably shorter than the entire sequence. Oligonucleotides can be used, for example, but should be at least about 10, preferably at least about 15, more preferably at least 20 nucleotides in length. When shorter length regions are used for comparison, a higher degree of sequence identity is required than for longer sequences. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (*See*, Gould, *et al.*, *PNAS USA* (1989) *86*:1934-1938.)

When longer nucleic acid fragments (>100 bp) are employed as probes, especially when using complete or large cDNA sequences, one can still screen with

moderately high stringencies (for example using 50% formamide at 37°C with minimal washing) in order to obtain signal from the target sample with 20-50% deviation, i.e., homologous sequences. (For additional information regarding screening techniques see Beltz, *et al.*, *Meth. Enzymology* (1983) *100*:266-285).

Another aspect of the present invention relates to isolated FtsZ polynucleotides. The polynucleotide sequences of the present invention include isolated polynucleotides that encode the polypeptides of the invention having a deduced amino acid sequence selected from the group of sequences set forth in the Sequence Listing and to other polynucleotide sequences closely related to such sequences and variants thereof.

The invention provides a polynucleotide sequence identical over its entire length to each coding sequence as set forth in the Sequence Listing. The invention also provides the coding sequence for the mature polypeptide or a fragment thereof, as well as the coding sequence for the mature polypeptide or a fragment thereof in a reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, pro-, or prepro- protein sequence. The polynucleotide can also include non-coding sequences, including for example, but not limited to, non-coding 5' and 3' sequences, such as the transcribed, untranslated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence that encodes additional amino acids. For example, a marker sequence can be included to facilitate the purification of the fused polypeptide. Polynucleotides of the present invention also include polynucleotides comprising a structural gene and the naturally associated sequences that control gene expression.

The invention also includes polynucleotides of the formula:

$$X\text{-}(R_1)_n\text{-}(R_2)\text{-}(R_3)_n\text{-}Y$$

wherein, at the 5' end, X is hydrogen, and at the 3' end, Y is hydrogen or a metal, R_1 and R_3 are any nucleic acid residue, n is an integer between 1 and 3000, preferably between 1 and 1000 and R_2 is a nucleic acid sequence of the invention, particularly a nucleic acid sequence selected from the group set forth in the Sequence Listing and preferably SEQ ID NOs:1,3,5,7,8,and 10-31. In the formula, R_2 is oriented so that its 5' end residue is at the left, bound to R_1, and its 3' end residue is at the right, bound to

R_3. Any stretch of nucleic acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

The invention also relates to variants of the polynucleotides described herein that encode for variants of the polypeptides of the invention. Variants that are fragments of the polynucleotides of the invention can be used to synthesize full-length polynucleotides of the invention. Preferred embodiments are polynucleotides encoding polypeptide variants wherein 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues of a polypeptide sequence of the invention are substituted, added or deleted, in any combination. Particularly preferred are substitutions, additions, and deletions that are silent such that they do not alter the properties or activities of the polynucleotide or polypeptide.

Further preferred embodiments of the invention that are at least 50%, 60%, or 70% identical over their entire length to a polynucleotide encoding a polypeptide of the invention, and polynucleotides that are complementary to such polynucleotides. More preferable are polynucleotides that comprise a region that is at least 80% identical over its entire length to a polynucleotide encoding a polypeptide of the invention and polynucleotides that are complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length are particularly preferred, those at least 95% identical are especially preferred. Further, those with at least 97% identity are highly preferred and those with at least 98% and 99% identity are particularly highly preferred, with those at least 99% being the most highly preferred.

Preferred embodiments are polynucleotides that encode polypeptides that retain substantially the same biological function or activity as the mature polypeptides encoded by the polynucleotides set forth in the Sequence Listing.

The invention further relates to polynucleotides that hybridize to the above-described sequences. In particular, the invention relates to polynucleotides that hybridize under stringent conditions to the above-described polynucleotides. As used herein, the terms "stringent conditions" and "stringent hybridization conditions" mean that hybridization will generally occur if there is at least 95% and preferably at least 97% identity between the sequences. An example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising 50% formamide,

5x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/milliliter denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at approximately 65°C. Other hybridization and wash conditions are well known and are exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, cold Spring Harbor, NY (1989), particularly Chapter 11.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set for in the Sequence Listing under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers as described herein.

As discussed herein regarding polynucleotide assays of the invention, for example, polynucleotides of the invention can be used as a hybridization probe for RNA, cDNA, or genomic DNA to isolate full length cDNAs or genomic clones encoding a polypeptide and to isolate cDNA or genomic clones of other genes that have a high sequence similarity to a polynucleotide set forth in the Sequence Listing. Such probes will generally comprise at least 15 bases. Preferably such probes will have at least 30 bases and can have at least 50 bases. Particularly preferred probes will have between 30 bases and 50 bases, inclusive.

The coding region of each gene that comprises or is comprised by a polynucleotide sequence set forth in the Sequence Listing may be isolated by screening using a DNA sequence provided in the Sequence Listing to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to identify members of the library which hybridize to the probe. For example, synthetic oligonucleotides are prepared which correspond to the FtsZ EST sequences. The oligonucleotides are used as primers in polymerase chain reaction (PCR) techniques to obtain 5' and 3' terminal sequence of FtsZ genes. Alternatively, where oligonucleotides of low degeneracy can be prepared from particular FtsZ

peptides, such probes may be used directly to screen gene libraries for FtsZ gene sequences. In particular, screening of cDNA libraries in phage vectors is useful in such methods due to lower levels of background hybridization.

Typically, a FtsZ sequence obtainable from the use of nucleic acid probes will show 60-70% sequence identity between the target FtsZ sequence and the encoding sequence used as a probe. However, lengthy sequences with as little as 50-60% sequence identity may also be obtained. The nucleic acid probes may be a lengthy fragment of the nucleic acid sequence, or may also be a shorter, oligonucleotide probe. When longer nucleic acid fragments are employed as probes (greater than about 100 bp), one may screen at lower stringencies in order to obtain sequences from the target sample which have 20-50% deviation (i.e., 50-80% sequence homology) from the sequences used as probe. Oligonucleotide probes can be considerably shorter than the entire nucleic acid sequence encoding an FtsZ enzyme, but should be at least about 10, preferably at least about 15, and more preferably at least about 20 nucleotides. A higher degree of sequence identity is desired when shorter regions are used as opposed to longer regions. It may thus be desirable to identify regions of highly conserved amino acid sequence to design oligonucleotide probes for detecting and recovering other related FtsZ genes. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (*See*, Gould, *et al.*, *PNAS USA* (1989) *86*:1934-1938.).

Another aspect of the present invention relates to FtsZ polypeptides. Such polypeptides include isolated polypeptides set forth in the Sequence Listing, as well as polypeptides and fragments thereof, particularly those polypeptides which exhibit FtsZ activity and also those polypeptides which have at least 50%, 60% or 70% identity, preferably at least 80% identity, more preferably at least 90% identity, and most preferably at least 95% identity to a polypeptide sequence selected from the group of sequences set forth in the Sequence Listing, and also include portions of such polypeptides, wherein such portion of the polypeptide preferably includes at least 30 amino acids and more preferably includes at least 50 amino acids.

"Identity", as is well understood in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence

relatedness between polypeptide or polynucleotide sequences, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods including, but not limited to, those described in *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York (1988); *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M. and Griffin, H.G., eds., Humana Press, New Jersey (1994); *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press (1987); *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., Stockton Press, New York (1991); and Carillo, H., and Lipman, D., SIAM *J Applied Math*, 48:1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available programs. Computer programs which can be used to determine identity between two sequences include, but are not limited to, GCG (Devereux, J., et al., *Nucleic Acids Research* 12(1):387 (1984); suite of five BLAST programs, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology, 12:* 76-80 (1994); Birren, *et al., Genome Analysis, 1:* 543-559 (1997)). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH, Bethesda, MD 20894; Altschul, S., *et al., J. Mol. Biol.*, 215:403-410 (1990)). The well known Smith Waterman algorithm can also be used to determine identity.

Parameters for polypeptide sequence comparison typically include the following:

Algorithm: Needleman and Wunsch, *J. Mol. Biol.* 48:443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci USA* 89:10915-10919 (1992)

Gap Penalty: 12

Gap Length Penalty: 4

A program which can be used with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison Wisconsin. The above

parameters along with no penalty for end gap are the default parameters for peptide comparisons.

Parameters for polynucleotide sequence comparison include the following:

Algorithm: Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970)

Comparison matrix: matches = +10; mismatches = 0

Gap Penalty: 50

Gap Length Penalty: 3

A program which can be used with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison Wisconsin. The above parameters are the default parameters for nucleic acid comparisons.

The invention also includes polypeptides of the formula:

$$X\text{-}(R_1)_n\text{-}(R_2)\text{-}(R_3)_n\text{-}Y$$

wherein, at the amino terminus, X is hydrogen, and at the carboxyl terminus, Y is hydrogen or a metal, R_1 and R_3 are any amino acid residue, n is an integer between 1 and 1000, and R_2 is an amino acid sequence of the invention, particularly an amino acid sequence selected from the group set forth in the Sequence Listing and preferably SEQ ID NOs: 2,4,6, and 9. In the formula, R_2 is oriented so that its amino terminal residue is at the left, bound to R_1, and its carboxy terminal residue is at the right, bound to R_3. Any stretch of amino acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

Polypeptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising a sequence selected from the group of a sequence contained in the Sequence Listing set forth herein .

The polypeptides of the present invention can be mature protein or can be part of a fusion protein.

Fragments and variants of the polypeptides are also considered to be a part of the invention. A fragment is a variant polypeptide which has an amino acid sequence that is entirely the same as part but not all of the amino acid sequence of the previously described polypeptides. The fragments can be "free-standing" or comprised within a larger polypeptide of which the fragment forms a part or a region, most preferably as a single continuous region. Preferred fragments are biologically

active fragments which are those fragments that mediate activities of the polypeptides of the invention, including those with similar activity or improved activity or with a decreased activity. Also included are those fragments that antigenic or immunogenic in an animal, particularly a human.

Variants of the polypeptide also include polypeptides that vary from the sequences set forth in the Sequence Listing by conservative amino acid substitutions, substitution of a residue by another with like characteristics. In general, such substitutions are among Ala, Val, Leu and Ile; between Ser and Thr; between Asp and Glu; between Asn and Gln; between Lys and Arg; or between Phe and Tyr. Particularly preferred are variants in which 5 to 10; 1 to 5; 1 to 3 or one amino acid(s) are substituted, deleted, or added, in any combination.

Variants that are fragments of the polypeptides of the invention can be used to produce the corresponding full length polypeptide by peptide synthesis. Therefore, these variants can be used as intermediates for producing the full-length polypeptides of the invention.

The polynucleotides and polypeptides of the invention can be used, for example, in the transformation of host cells, such as plant host cells, as further discussed herein.

The invention also provides polynucleotides that encode a polypeptide that is a mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids within the mature polypeptide (for example, when the mature form of the protein has more than one polypeptide chain). Such sequences can, for example, play a role in the processing of a protein from a precursor to a mature form, allow protein transport, shorten or lengthen protein half-life, or facilitate manipulation of the protein in assays or production. It is contemplated that cellular enzymes can be used to remove any additional amino acids from the mature protein.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. The inactive precursors generally are activated when the prosequences are removed. Some or all of the prosequences may be removed prior to activation. Such precursor protein are generally called proproteins.

Homologous sequences are found when there is an identity of sequence and can be determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions between a known FtsZ and a candidate source. Conservative changes, such as Glu/Asp, Val/Ile, Ser/Thr, Arg/Lys and Gln/Asn can also be considered in determining sequence homology. Typically, a lengthy nucleic acid sequence can show as little as 50-60% sequence identity, and more preferably at least about 70% sequence identity, between the target sequence and the given FtsZ sequence of interest excluding any deletions which can be present, and still be considered related. Amino acid sequences are considered homologous by as little as 25% sequence identity between the two complete mature proteins. (*See* generally, Doolittle, R.F., *OF URFS and ORFS* (University Science Books, CA, 1986.)

In addition, not only can sequences provided herein be used to identify homologous FtsZ sequences, but the resulting sequences obtained therefrom can also provide a further method to obtain FtsZ sequences from other plant and/or bacterial sources. In particular, PCR can be a useful technique to obtain related FtsZ sequences from sequence data provided herein. One skilled in the art will be able to design oligonucleotide probes based upon sequence comparisons or regions of typically highly conserved sequence.

Once the nucleic acid sequence is obtained, the transcription, or transcription and translation (expression), of the FtsZ sequence in a host cell is desired to produce a ready source of the enzyme and/or modify the number and/or size of the plastids found therein. Other useful applications can be found when the host cell is a plant host cell, *in vitro* and *in vivo*.

Nucleic acids (genomic DNA, plasmid DNA, cDNA, synthetic DNA, mRNA, etc.) encoding FtsZ or amino acid sequences of the purified enzymes, which permit design of nucleic acid probes facilitating the isolation of DNA coding sequences therefor, are known in the art and are available for use in the methods of the present invention. It is generally recognized to an artisan skilled in the field to which the present invention pertains that the nucleic acid sequences provided herein and the amino acid sequences derived therefrom can be used to isolate other potential FtsZ genes from GenBank using DNA and peptide search techniques generally known in the art.

In addition to the sequences described in the present invention, DNA coding sequences useful in the present invention can be derived from algae, fungi, bacteria, plants, etc. Homology searches in existing databases using signature sequences corresponding to conserved nucleotide and amino acid sequences of FtsZ can be employed to isolate equivalent, related genes from other sources such as plants and microorganisms. Searches in EST databases can also be employed. Furthermore, the use of DNA sequences encoding enzymes functionally enzymatically equivalent to those disclosed herein, wherein such DNA sequences are degenerate equivalents of the nucleic acid sequences disclosed herein in accordance with the degeneracy of the genetic code, is also encompassed by the present invention. Demonstration of the functionality of coding sequences identified by any of these methods can be carried out by complementation of mutants of appropriate organisms, such as *E. coli*. The sequences of the DNA coding regions can be optimized by gene resynthesis, based on codon usage, for maximum expression in particular hosts.

The nucleic acid sequences which encode FtsZ can be used in various constructs, for example, as probes to obtain further sequences. Alternatively, these sequences can be used in conjunction with appropriate regulatory sequences to increase levels of the respective FtsZ sequence of interest in a host cell for recovery or study of the enzyme *in vitro* or *in vivo* or to decrease levels of the respective FtsZ sequence of interest for some applications when the host cell is a plant entity, including plant cells, plant parts (including but not limited to seeds, cuttings or tissues) and plants.

Thus, depending upon the intended use, the constructs can contain the nucleic acid sequence which encodes the entire FtsZ protein, or a portion thereof. For example, where antisense inhibition of a given FtsZ protein is desired, the entire FtsZ sequence is not required. Furthermore, where FtsZ constructs are intended for use as probes, it can be advantageous to prepare constructs containing only a particular portion of a FtsZ encoding sequence, for example a sequence which is discovered to encode a highly conserved FtsZ region.

As discussed above, nucleic acid sequence encoding a plant or other FtsZ proteins of this invention can include genomic, cDNA or mRNA sequence. By "encoding" is meant that the sequence corresponds to a particular amino acid sequence

either in a sense or anti-sense orientation. By "extrachromosomal" is meant that the sequence is outside of the plant genome of which it is naturally associated. By "recombinant" is meant that the sequence contains a genetically engineered modification through manipulation via mutagenesis, restriction enzymes, and the like.

A cDNA sequence may or may not contain pre-processing sequences, such as transit peptide sequences or targeting sequences to facilitate delivery of the FtsZ protein to a given organelle or membrane location. The use of any such precursor FtsZ DNA sequence is preferred for uses in plant cell expression. A genomic FtsZ sequence can contain the transcription and translation initiation regions, introns, and/or transcript termination regions of the plant FtsZ, which sequences can be used in a variety of DNA constructs, with or without the FtsZ structural gene. Thus, nucleic acid sequences corresponding to the FtsZ sequences of this invention can also provide signal sequences useful to direct protein delivery into a particular organellar or membrane location, 5' upstream non-coding regulatory regions (promoters) having useful tissue and timing profiles, 3' downstream non-coding regulatory regions useful as transcriptional and translational regulatory regions, and may lend insight into other features of the gene.

Once the desired plant or other FtsZ nucleic acid sequence is obtained, it can be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions can be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions can be performed on the naturally occurring sequence. In addition, all or part of the sequence can be synthesized. In the structural gene, one or more codons can be modified to provide for a modified amino acid sequence, or one or more codon mutations can be introduced to provide for a convenient restriction site or other purpose involved with construction or expression. The structural gene can be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

For the most part, the constructs will involve regulatory regions functional in plants which provide for altered size and number of plastids in a plant cell. The open reading frame, coding for the FtsZ protein, FtsZ-related protein or functional fragment thereof will be joined at its 5' end to a transcription initiation regulatory region such as

the wild-type sequence naturally found 5' upstream to theFtsZ or FtsZ-related structural gene, or to a heterologous regulatory region from a gene naturally expressed in plant tissues. Examples of useful plant regulatory gene regions include those from T-DNA genes, such as nopaline or octopine synthase, plant virus genes, such as CaMV 35S, or from native plant genes.

The DNA sequence encoding a plant or other FtsZ protein of this invention can be employed in conjunction with all or part of the gene sequences normally associated with the FtsZ. In its component parts, a DNA sequence encoding FtsZ is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription and translation in a host cell, the DNA sequence encoding plant FtsZ and a transcription and translation termination region.

Potential host cells include both prokaryotic and eukaryotic cells. A host cell can be unicellular or found in a multicellar differentiated or undifferentiated organism depending upon the intended use. Cells of this invention can be distinguished by having a FtsZ sequence foreign to the wild-type cell present therein, for example, by having a recombinant nucleic acid construct encoding a FtsZ protein therein not native to the host species.

Depending upon the host, the regulatory regions will vary, including regions from viral, plasmid or chromosomal genes, or the like. For expression in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters can be employed. Expression in a microorganism can provide a ready source of the plant enzyme. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli, B. subtilis, Sacchromyces cerevisiae,* including genes such as beta-galactosidase, T7 polymerase, tryptophan E and the like.

In a preferred embodiment, the constructs will involve regulatory regions functional in plants which provide for modified production of plant FtsZ, and, possibly, modification of the plant cell plastid. The open reading frame coding for the plant FtsZ or functional fragment thereof will be joined at its 5' end to a transcription initiation regulatory region. In embodiments wherein the expression of the FtsZ protein is desired in a plant host, the use of all or part of the complete plant FtsZ gene

is desired; namely all or part of the 5' upstream non-coding regions (promoter) together with the structural gene sequence and 3' downstream non-coding regionscan be employed.

If a different promoter is desired, such as a promoter native to the plant host of interest or a modified promoter, i.e., having transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source, numerous transcription initiation regions are available which provide for a wide variety of constitutive or regulatable, e.g., inducible, transcription of the structural gene functions. The transcription/translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. Among transcriptional initiation regions used for plants are such regions associated with the T-DNA structural genes such as for nopaline and mannopine synthases, the 19S and 35S promoters from CaMV, and the 5' upstream regions from other plant genes such as napin, ACP, SSU, PG, zein, phaseolin E, and the like. Enhanced promoters, such as double 35S, are also available for expression of FtsZ sequences. For such applications when 5' upstream non-coding regions are obtained from other genes regulated during seed maturation, those preferentially expressed in plant embryo tissue, such as ACP and napin-derived transcription initiation control regions, are desired. Such "seed-specific promoters" can be obtained and used in accordance with the teachings of issued U.S. Patent Numbers 5,608,152 and 5,530,194, which references are hereby incorporated by reference. Transcription initiation regions which are preferentially expressed in seed tissue, i.e., which are undetectable in other plant parts, are considered desirable for TAG modifications in order to minimize any disruptive or adverse effects of the gene product.

Regulatory transcript termination regions can be provided in DNA constructs of this invention as well. Transcript termination regions can be provided by the DNA sequence encoding the plant FtsZ or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally associated with the transcript initiation region. Where the transcript termination region is from a different gene source, it will contain at least about 0.25 kb, preferably about 1-3 kb of sequence 3' to the structural gene from which the termination region is derived.

Plant expression or transcription constructs having a plant FtsZ as the DNA sequence of interest for increased or decreased expression thereof can be employed with a wide variety of plant life, particularly, plant life involved in the production of vegetable oils for edible and industrial uses. Most especially preferred are temperate oilseed crops. Plants of interest include, but are not limited to, rapeseed (Canola and High Erucic Acid varieties), sunflower, safflower, cotton, soybean, peanut, coconut and oil palms, and corn. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences can be required. Importantly, this invention is applicable to dicotyledenous and monocotyledenous species alike and will be readily applicable to new and/or improved transformation and regulation techniques.

The method of transformation is not critical to the instant invention; various methods of plant transformation are currently available. As newer methods are available to transform crops, they can be directly applied hereunder. For example, many plant species naturally susceptible to *Agrobacterium* infection can be successfully transformed via tripartite or binary vector methods of *Agrobacterium*-mediated transformation. In addition, techniques of microinjection, DNA particle bombardment, and electroporation have been developed which allow for the transformation of various monocot and dicot plant species.

In developing the DNA construct, the various components of the construct or fragments thereof will normally be inserted into a convenient cloning vector which is capable of replication in a bacterial host, e.g., *E. coli*. Numerous vectors exist that have been described in the literature. After each cloning, the plasmid can be isolated and subjected to further manipulation, such as restriction, insertion of new fragments, ligation, deletion, insertion, resection, etc., so as to tailor the components of the desired sequence. Once the construct has been completed, it can then be transferred to an appropriate vector for further manipulation in accordance with the manner of transformation of the host cell.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene can provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an

auxotrophic host, viral immunity or the like. Depending upon the number of different host species in which the expression construct or components thereof are introduced, one or more markers can be employed, where different conditions for selection are used for the different hosts. A number of markers have been developed for use for selection of transformed plant cells, such as those which provide resistance to various antibiotics, herbicides, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

As mentioned above, the manner in which the DNA construct is introduced into the plant host is not critical to this invention. Any method which provides for efficient transformation can be employed. Various methods for plant cell transformation include the use of Ti- or Ri-plasmids, microinjection, electroporation, DNA particle bombardment, liposome fusion, or the like. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders can find use with other modes of transformation.

Once a transgenic plant is obtained which contains cells with altered numbers and/or sizes of chloroplasts, tissue containing such cells can then be used in plastid transformation experiments. For example, utilizing tissue containing cells with larger plastids provides for a larger target in plastid transformation methods, thus allowing for an increased probability of introduction of the foreign DNA into the plant cell plastid.

The DNA sequences, or polynucleotides, for use in plastid transformation of this invention will contain a plastid expression construct generally comprising a promoter functional in a plant cell plastid, and a DNA sequence of interest to be expressed in the transformed plastid cells.

Constructs and methods for use in transforming the plastids of higher plants are described in Zoubenko *et al.* (*Nuc Acid Res* (1994) 22(19):3819-3824), Svab *et al.* (*Proc. Natl. Acad. Sci.*(1990) 87:8526-8530 and *Proc. Natl. Acad. Sci.*(1993) 90:913-917) and Staub *et al.* (*EMBO J.* (1993) 12:601-606). Constructs and methods for use

in transforming plastids of higher plants to express DNA sequences under the control of a nuclearly encoded, plastid targeted T7 polymerase are described in U.S. Patent Number 5,576,198. The complete DNA sequences of the plastid genome of tobacco are reported by Shinozaki *et al.* (*EMBO J.* (1986) 5:2043-2049).

Stable transformation of tobacco plastid genomes by particle bombardment is reportcd (Svab *et.al.* (1990), *supra*) and Svab *et al.* (1993), *supra*). The methods described therein can be employed to obtain plants homoplasmic for plastid expression constructs using the methods described herein. Briefly, such methods involve DNA bombardment of a target host explant, preferably from a tissue which is rich in metabolically active plastid organelles, such as green plant tissues including leaves, and cotyledons. The bombarded tissue is then cultured for ~2 days on a cell division promoting media. The plant tissue is then transferred to a selective media containing an inhibitory amount of the particular selective agent, as well as the particular hormones and other substances necessary to obtain regeneration for that particular plant species. For example, in the above publications and the examples provided herein, the selective marker is the bacterial *aad*A gene and the selective agent is spectinomycin. The *aad*A gene product allows for continued growth and greening of cells whose chloroplasts comprise the marker gene product. Cells which do not contain the marker gene product are bleached. The bombarded explants will form green shoots in approximately 3-8 weeks. Leaves from these shoots are then subcultured on the same selective media to ensure production and selection of homoplasmic shoots. As an alternative to a second round of shoot formation, the initial selected shoots can be grown to mature plants and segregation relied upon to provide transformed plants homoplastic for the inserted gene construct.

The transformed plants so selected can then be analyzed to determine whether the entire plastid content of the plant has been transformed (homoplastic transformants). Typically, following two rounds of shoot formation and spectinomycin selection, approximately 50% of the transgenic plantlets analyzed are homoplastic as determined by Southern blot analysis of plastid DNA. These plantlets are selected for further cultivation, both for analysis of the transgenic plastid phenotype (where the nuclear viral polymerase expression construct is also present in

the plastid transformant), or for use in methods to transform the viral polymerase construct into the nucleus of the transplastomic plants.

The methods of the present invention provide for a more efficient approach to obtaining homoplasmic plants. Wild-type plant cells typically contain 50 to 100 plastids per cell. However, once a transplastomic plant is obtained, the DNA sequence contained in the plant cell nucleus can be crossed away from the transplastomic cells. The DNA sequence transformed into the nucleus encoding for the alteration can be crossed away from the plant containing the transformed plastids. Once the DNA sequence has been crossed out, the plastids in the host plant cell can divide and revert back to normal (i.e. wild-type) plastid size and numbers. By applying the selective agent for which the plastid expression constructs provides resistance, cells containing a pure population of the plastids containing the foreign DNA can be obtained.

The vectors for use in plastid transformation preferably include means for providing a stable transfer of the plastid expression construct and selectable marker construct into the plastid genome. This is most conveniently provided by regions of homology to the target plastid genome. The regions of homology flank the construct to be transferred and provide for transfer to the plastid genome by homologous recombination, via a double crossover into the genome. The complete DNA sequence of the plastid genome of tobacco has been reported (Shinozaki *et al.*, *EMBO J.* (1986) *5*:2043-2049). Complete DNA sequences of the plastid genomes from liverwort (Ohyama *et al.*, *Nature* (1986) *322*:572-574) and rice (Hiratsuka *et al.*, *Mol. Gen. Genet.* (1989) *217*:185-194), have also been reported.

Where the regions of homology are present in the inverted repeat regions of the plastid genome (known as IRA and IRB), two copies of the transgene are expected per transformed plastid. Where the regions of homology are present outside the inverted repeat regions of the plastid genome, one copy of the transgene is expected per transformed plastid. The regions of homology within the plastid genome are approximately 1kb in size. Smaller regions of homology can also be used, and as little as 100 bp can provide for homologous recombination into the plastid genome. However, the frequency of recombination and thus the frequency of obtaining plants having transformed plastids decreases with decreasing size of the homology regions.

Examples of constructs comprising such regions of homology for tobacco plastid transformation are described in Svab *et.al.* (1990 supra) and Svab and Maliga (1993 supra). Regions useful for recombination into tobacco and *Brassica* plastid genomes are also described in the following examples. Similar homologous recombination and selection constructs can be prepared using plastid DNA from the target plant species.

Other means of transfer to the plastid genome are also considered herein, such as by methods involving the use of transposable elements. For example, the constructs to be transferred into the plastid genome can be flanked by the inverted repeat regions from a transposable marker which functions in plant plastids. A DNA construct which provides for transient expression of the transposase required to transfer the target DNA into the plastids is also introduced into the chloroplasts. In this manner, a variety of phenotypes can be obtained in plants transformed with the same expression construct depending on positional effects which can result from insertion of the expression constructs into various locations on the plastid genome. Appropriate transposons for use in such plastic transformation methods include bacterial Tn10, bacteriophage Mu and various other known bacterial transposons.

The DNA sequence of interest in the plastid promoter expression constructs can be an encoding sequence which is oriented for expression of a particular structural gene, such that the protein encoded by the structural gene sequence is produced in the transformed plastid. In addition, the DNA sequence of interest can include a number of individual structural gene encoding regions such that an operon for expression of a number of genes from a single plastid promoter region is produced. Thus, it is possible to introduce and express multiple genes from an engineered or synthetic operon or from a pre-existing prokaryotic gene cluster. Such a method would allow large scale and inexpensive production of valuable proteins and fine chemicals in a particular desired plant tissue or a particular stage of development, depending upon the promoter used to drive nuclear expression of the specific viral polymerase. Such an approach is not practical by standard nuclear transformation methods since each gene must be engineered into a monocistron including an encoded transit peptide for plastid uptake and appropriate promoter and terminator signals. As a result, gene expression levels would be expected to vary widely between cistrons, and generation of a number of transgenic plant lines would be required. Ultimately crosses would be

required to introduce all of these cistrons into one plant to get expression to the target biochemical pathway.

Alternatively, the DNA sequence of interest in the plastid construct can be a fragment of an endogenous plastid gene oriented such that an RNA complementary to the endogenous gene mRNA is produced in the transformed plastid. Such antisense constructs can be used to decrease the expression of the target plastid gene.

In order to provide a means of selecting the desired plant cells following plastid transformation, the polynucleotides for plastid transformation will also contain a construct which provides for expression of a marker gene. Expression of the marker gene product allows for selection of plant cells comprising plastid organelles which are expressing the marker protein. In the examples provided herein, a bacterial *aad*A gene is expressed under the regulatory control of chloroplast 5' promoter and 3' transcription termination regions. The use of such an expression construct for plastid transformation of plant cells has been described by Svab and Maliga (1993, *supra*). Expression of the *aad*A gene confers resistance to spectinomycin and streptomycin, and thus allows for the identification of plant cells expressing this marker gene. Selection for the *aad*A marker gene is based on identification of plant cells which are not bleached by the presence of streptomycin, or more preferably spectinomycin, in the plant growth medium. Other genes which encode a product involved in chloroplast metabolism can also be used as selectable markers. For example, genes which provide resistance to plant herbicides such as glyphosate, bromoxynil or imidazolinone can find particular use. Such genes have been reported by Stalker *et al.* (*J. Biol. Chem.* (1985) *260*:4724-4728; glyphosate resistant EPSP), Stalker *et al.* (*J. Biol. Chem.* (1985) *263*:6310-6314; bromoxynil resistant nitrilase gene), and Sathasivan *et al.* (*Nucl. Acids Res.* (1990) *18*:2188; AHAS imidazolinone resistance gene).

The present invention also provides methods for obtaining a plastid transformed plant on medium containing glyphosate. At the initial event of transformation only a few plastids out of the many present in a plant cell are transformed and therefore are able to express glyphosate resistant marker gene product. The rest of the untransformed plastids within the cell remains vulnerable to the effect of glyphosate. Therefore, although the cell contains transformed plastids, it

is unable to divide and sort out the transformed plastid resulting in lack of recovery of transformed callus tissue which would give rise to the transformed regenerants. Thus, any method that reduces plastid number to one or few within the cell has the potential to survive the effect of glyphosate and be useful as selectable marker for plastid transformation.

The following examples are provided by way of illustration and not by way of limitation.

EXAMPLES

Example 1: Identification of Plant *ftsZ* Sequences

In order to obtain a plant tissue source with an altered number and/or size of plastids using antisense and/or sense expression of the bacterial FtsZ plant homologues, public as well as proprietary sequence databases are queried for homologous sequences in soybean, rice, *Arabidsopsis*, corn and *Brassica*. Two types of plant FtsZ proteins have been previously identified in GenBank, type I FtsZ proteins exemplified by accession gil1079731 (SEQ ID NO:32), appear to be imported into the plastid, while type II FtsZ proteins, exemplified by accession gil3608494 (SEQ ID NO:33) and gil683524 (SEQ ID NO:34), appear to remain in the cytoplasm. Homologs of both the type I FtsZ sequence as well as homologues of type II FtsZ genes are described below. The sequences used to search against the databases are: type I FtsZ homologue search was (SEQ ID NO:32), and for type II FtsZ searches, (SEQ ID NO:33) is used.

Searches performed in proprietary databases containing sequences obtained from *Arabidopsis* identified DNA sequences which are related to the FtsZ1 sequence. The sequence of SEQ ID NO:1 is identified as AtFtsZ1. The deduced amino acid sequence encoded by SEQ ID NO:1 is provided in SEQ ID NO:2. In addition, one sequence (SEQ ID NO:3) was identified as related to the FtsZ2 sequence. The deduced amino acid sequence encoded by SEQ ID NO:3 is provided in SEQ ID NO:4.

Sequences were also identified in databases containing sequences obtained from *Brassica*. One sequence was identified as related to the *Arabidopsis* FtsZ1 sequence. Based on sequence alignments between the two sequences, approximately 170 amino acids were predicted to be missing from the *Brassica* sequence at the N-terminus. To obtain a full length coding sequence for the *Brassica* FtsZ1 (BnFtsZ1) gene, RACE PCR using DNA obtained from *Brassica* leaves was performed using the primers SC258 (SEQ ID NO:35) and SC259 (SEQ ID NO:36). One reaction product was found to contain the most 5' sequence (SEQ ID NO:70) and was used to produce a full length sequence referred to as BnFtsZ1 (SEQ ID NO:5). The deduced amino acid sequence encoded by BnFtsZ1 is provided in SEQ ID NO:6)

A FtsZ1 homolog was also identified in tobacco with PCR using primers designed to the conserved amino acid domains of the *Arabidopsis* FtsZ1 sequence. The PCR primers used are identified as SC252 (SEQ ID NO:37), SC253 (SEQ ID NO:38), SC254 (SEQ ID NO:39) and SC255 (SEQ ID NO:40). The reaction products were cloned into TOPO TA (Invitrogen), and a single clone, referred to as xanthil-26-contig (SEQ ID NO:7), contained the most sequence. Additional primers were designed for use in RACE PCR to obtain full length coding sequence for the tobacco FtsZ1 homolog. For amplification of the 5' region, primers SC291 (SEQ ID NO:41) and SC292 (SEQ ID NO:42) were used, and for amplification of the 3' sequence, primers SC293 (SEQ ID NO:43) and SC294 (SEQ ID NO:44) were used. The PCR products were cloned in TOPO TA and sequenced. Clone xanftsZ1-5'-15 (SEQ ID NO:71) was chosen to be the best for the 5' tobacco FtsZ1 sequence since it contained the greatest amount of 5' sequence and overlap with xanthil-26-contig. This sequence was combined with the xanthil-26-contig to produce xanFtsZ1 (SEQ ID NO:8). The deduced amino acid sequence is provided in SEQ ID NO:9.

FtsZ homolog sequences were identified in databases containing DNA sequences obtained from corn by BLAST searches using the Arabidopsis FtsZ1 and FtsZ2 amino acid sequences. Ten sequence were identified as related to these FtsZ sequences, provided in SEQ ID NOs:10-19. The clones, when aligned, revealed six contigs, and the best representative clone for each were chosen for further analysis. Sequence analysis of SEQ ID NO:10 revealed a high homology to AtFtsZ1, and was estimated to be missing 158 amino acids at the N-terminal end when compared to

Arabidopsis FtsZ1. Clone SEQ ID NO:13 was found to overlap perfectly with SEQ ID NO:10 for 153 nt at the 5'end and in addition had 167 nt additional nt at the 5' end that had amino acid homology with the *Arabidopsis* FtsZ1. However, this clone was also not predicted to encode the full-length FtsZ, and was still missing 113 amino acids at the N-terminal end when compared to Arabidopsis FtsZ1. Interestingly, for clone SEQ ID NO:13, its homology with SEQ ID NO:10, ends at position 167nt and diverges. This could either be indicative of the presence of intronic sequence or a new class of FtsZ protein. Primer SC321 (SEQ ID NO:45) was designed to pull out the missing maize FtsZ1 sequence by RACE PCR.

Sequence analysis of SEQ ID NO:18 revealed its high homology to FtsZ2, and was also predicted to not to be full-length and missing about 286 amino acids at the N-terminal end when compared to Arabidopsis FtsZ2. Primer SC322 (SEQ ID NO:46) was designed to pull out the missing maize FtsZ2 sequence by RACE PCR. Although SEQ ID NO;14 and SEQ ID NO:15 were identified with the highest BLAST scores with FtsZ2.

Soybean FtsZ homolog sequences were identified in databases by BLAST searches with Arabidopsis FtsZ1 and FtsZ2 amino acid sequences. Twelve sequences were obtained, and are provided in SEQ ID NOs:20-31. Sequence analysis of SEQ ID NO:20, SEQ ID NO:24 and SEQ ID NO:25 revealed high homology to FtsZ1 and none to be full-length when compared to Arabidopsis FtsZ1. SEQ ID NO:25 had the longest sequence at the N-terminal end and is predicted to be missing 64 amino acids at the N-terminal when compared to Arabidopsis FtsZ1 sequence. Sequences of SEQ ID NO:20, SEQ ID NO:24 and SEQ ID NO:25 were used to correct the overlapping region. RACE PCR primers can now be designed to amplify the ends for obtaining a full length DNA sequence.

A sequence alignment between the Arabidopsis, Brassica, tobacco, soybean, and corn FtsZ1 protein sequences is provided in figure 1.

Example 2: Preparation of Plant Expression Constructs

2A. Nuclear Expression Constructs

Constructs are prepared for transformation into a plant cell nucleus for alteration of the plastid size and/or number in the transformed plant cell. Constructs can be prepared to alter the plastids constitutively, or in a tissue specific manner, for example, in leaf tissue, or seed tissue.

A plasmid containing the napin cassette derived from pCGN3223 (described in USPN 5,639,790, the entirety of which is incorporated herein by reference) was modified to make it more useful for cloning large DNA fragments containing multiple restriction sites, and to allow the cloning of multiple napin fusion genes into plant binary transformation vectors. An adapter comprised of the self annealed oligonucleotide of sequence CGCGATTTAAATGGCGCGCCCTGCAGGCGGCCGCCTGCAGGGCGCGCCAT TTAAAT (SEQ ID NO:47) was ligated into the cloning vector pBC SK+ (Stratagene) after digestion with the restriction endonuclease BssHII to construct vector pCGN7765. Plamids pCGN3223 and pCGN7765 were digested with NotI and ligated together. The resultant vector, pCGN7770, contains the pCGN7765 backbone with the napin seed specific expression cassette from pCGN3223.

The cloning cassette, pCGN7787, essentially the same regulatory elements as pCGN7770, with the exception of the napin regulatory regions of pCGN7770 have been replaced with the double CAMV 35S promoter and the tml polyadenylation and transcriptional termination region.

A binary vector for plant transformation, pCGN5139, was constructed from pCGN1558 (McBride and Summerfelt, (1990) Plant Molecular Biology, 14:269-276). The polylinker of pCGN1558 was replaced as a HindIII/Asp718 fragment with a polylinker containing unique restriction endonuclease sites, AscI, PacI, XbaI, SwaI, BamHI,and NotI. The Asp718 and HindIII restriction endonuclease sites are retained in pCGN5139.

A series of turbo binary vectors are constructed to allow for the rapid cloning of DNA sequences into binary vectors containing transcriptional initiation regions (promoters) and transcriptional termination regions.

The plasmid pCGN8618 was constructed by ligating oligonucleotides 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:48) and 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:49) into

SalI/XhoI-digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3' region was excised from pCGN8618 by digestion with Asp718I; the fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8622.

The plasmid pCGN8619 was constructed by ligating oligonucleotides 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC -3' (SEQ ID NO:50) and 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:51) into SalI/XhoI-digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3' region was removed from pCGN8619 by digestion with Asp718I; the fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8623.

The plasmid pCGN8620 was constructed by ligating oligonucleotides 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGGAGCT -3' (SEQ ID NO:52) and 5'-CCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:53) into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region was removed from pCGN8620 by complete digestion with Asp718I and partial digestion with NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert

orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8624.

The plasmid pCGN8621 was constructed by ligating oligonucleotides 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCCAGCT -3' (SEQ ID NO:54) and 5'-GGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:55) into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region was removed from pCGN8621 by complete digestion with Asp718I and partial digestion with NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8625.

The plasmid construct pCGN8640 is a modification of pCGN8624 described above. A 938bp PstI fragment isolated from transposon Tn7 which encodes bacterial spectinomycin and streptomycin resistance (Fling et al. (1985), *Nucleic Acids Research* 13(19):7095-7106), a determinant for E. coli and Agrobacterium selection, was blunt ended with Pfu polymerase. The blunt ended fragment was ligated into pCGN8624 that had been digested with SpeI and blunt ended with Pfu polymerase. The region containing the PstI fragment was sequenced to confirm both the insert orientation and the integrity of cloning junctions.

The spectinomycin resistance marker was introduced into pCGN8622 and pCGN8623 as follows. A 7.7 Kbp AvrII-SnaBI fragment from pCGN8640 was ligated to a 10.9 Kbp AvrII-SnaBI fragment from pCGN8623 or pCGN8622, described above. The resulting plasmids were pCGN8641 and pCGN8643, respectively.

The Arabidopsis FtsZ1 nucleotide sequence was used to construct the sense expression vector pCGN6495 for use in transformation of *Arabidopsis*, *Brassica* and tobacco. For this construct, the Arabidopsis ftsZ1 sequence was PCR amplified. To monitor protein expression of FtsZ1 in transformed lines, a c-myc tag (EQKLISEEDL

(SEQ ID NO:56)), was translationally fused to FtsZ1 at the C-terminal end. The PCR amplification was done by first round of amplification with primers SC247 (SEQ ID NO:57) and SC260 (SEQ ID NO:58) followed by amplification with SC247 (SEQ ID NO:59) and SC261 (SEQ ID NO:60) using the product of the first amplification as the template DNA, using standard amplification parameters. The final amplification product, FtsZ1/c-myc fusion was cloned in the nuclear transformation vector pCGN8624 to create pCGN6495, which was used to nuclear transform Arabidopsis, canola and tobacco using standard protocols.

The turbo vector pCGN8624 was used for the antisense constructs such that the antisense sequence is driven from d35S promoter. For *Arabidopsis* the coding sequence (from ATG to TAG) was amplified with primers SC248 (SEQ ID NO:61) and SC250 (SEQ ID NO:62) using AtFtsZ1 as template. For *Brassica*, primers SC276 (SEQ ID NO:63) and SC268 (SEQ ID NO:64) were used with PCR fragment SC3-1-1 (SEQ ID NO:70) as template DNA to generate a *Hind*III/*Pst*I fragment and cloned in pBSKS (Stratagene) to generate pCGN6528. Primer SC276 was designed to be located 140 bases downstream from ATG due to the presence of nonhomologous stretch of sequence compared to *Arabidopsis* FtsZ1 contained in the first 140 bases sequence fragment. The 3' half of the coding sequence was PCR amplified using primers SC269 (SEQ ID NO:65) and SC270 (SEQ ID NO:66) to produce a PstI/NotI fragment, and subsequently cloned in pCGN6528 to generate pCGN6529. The *Hind*III/*Pst*I fragment containing BnFtsZ1 sequence (from 140b downstream of ATG to TAG) was cloned in turbo vector pCGN8624 to generate final transformation vectors pCGN6530 and pCGN6611. The *Hind*III/*Not*I fragment containing BnFtsZ1 sequence was also cloned into pCGN8643 vector for seed-specific antisense FtsZ1 expression. For tobacco, primers SC305 and SC306 were designed to PCR amplify FtsZ1 sequence to produce a *Sse*I/*Not*I fragment using 5' RACE PCR library DNA made from leaf RNA, and cloned into TOPO TA2.1 to produce pCGN6565. The *Sse*I/*Not*I fragment from pCGN6565 was cloned in the turbo vector pCGN8624 to generate final transformation vector pCGN6566.

2A. Plastid Expression Constructs

Constructs and methods for use in transforming the plastids of higher plants are described in Zoubenko *et al.* (*Nuc Acid Res* (1994) 22(19):3819-3824), Svab *et al.* (*Proc. Natl. Acad. Sci.*(1990) 87:8526-8530 and *Proc. Natl. Acad. Sci.*(1993) 90:913-917) and Staub *et al.* (*EMBO J.* (1993) 12:601-606). Constucts and methods for use in transforming plastids of higher plants to express DNA sequences under the control of a nuclearly encoded, plastid targeted T7 polymerase are described in U.S. Patent Number 5,576,198. The complete DNA sequences of the plastid genome of tobacco are reported by Shinozaki *et al.* (*EMBO J.* (1986) 5:2043-2049).

A plastid expression construct, pMON49218, was constructed to express the synthetic CP4 sequence with the 14 amino acid GFP fusion from the promoter region of the 16SrDNA operon having the nuclear-encoded RNA polymerase region (PrrnPEP+NEP), and the terminator region from the plastid rps16 gene. The DNA sequence of the Prrn/NEP/G10L::14aaGFP fusion SEQ ID NO:67.

Example 3: Plant Transformation And Analysis

Constructs for the expression of sense or antisense sequences are transformed into tobacco cells using the methods described by Ursin *et. al.* (1991) *Plant Cell* 3:583-591.

Transgenic tobacco plants containing the nuclear FtsZ constructs were analyzed for alterations in plastid morphology, including size and number of plastids present in the plant cell.

Fifty-eight initial transformants (T1 generation) obtained from transformation with FtsZ1 expression construct pCGN6495 were screened for the large plastid phenotype and divided into three categories. Thirty-four (34) lines contained less than 5 large plastids, 8 lines contained between 5-20 plastids and 16 lines more than 20 (wild-type# and more than wild-type#) plastids. One line, Nt6495-61, contained a single large plastid.

The screening method involved examining isolated mesophyll protoplasts at 100X magnification under light microscope. The large plastid containing transgenic

plants appear to be phenotypically indistinguishable from wild-type under culture and greenhouse conditions.

Estimation of plastid DNA copy number from several large plastid lines revealed no difference when compared to wild-type. Southern analysis was used to estimate transgene copy number in the large plastid lines and several lines with single integration events were identified. Western analysis of the large plastid lines with c-myc antibody confirmed expression of the introduced transgene (tagged by c-myc). T2 seeds were collected from selected plants from each of the three categories.

Example 4: Plastid Transformation and Analysis

Leaf material from three transgenic lines, Nt6495-30 (with <5plastids/cell), Nt6495-16 (with 5-20 plastids/cell) and Nt6495-69 (with 5-20 plastids/cell), were obtained for evaluation of plastid transformation efficiency and direct glyphosate selection. Plastid transformation vector pMON49218 which contains *aadA* gene for spectinomycin selection and GFP as a marker was used to bombard 15 leaf explants of each of the three transgenic lines. For each series of bombardment of the transgenic line 15 wild type control leaves were used. The order of bombardment for the transgenic line and the wild type leaves were randomized to eliminate any bias.

Transformation frequency of one event Nt6495-30 was approximately double that of the wild type control producing 7 versus 3 transformants respectively. Nt6495-16 and Nt6495-69 had approximately the same transformation frequency (3 transformants) as the control. Thus, our preliminary analysis reveals that plastid transformation efficiency can have been enhanced by reducing the plastid number from wild type to less than 5 plastids per cell. Interestingly, all of the plastid transformant regenerants from Nt6495 lines were very much slower in growth and size compared to those from wild type. It appears that the presence of the selectable antibiotic spectinomycin dihydrochloride at a concentration of 500mg/ml can have affected the regenerability of cells in the Nt6495 lines. Thus, it is possible that there could be more plastid transformed cells in the transgenic Nt6495 lines which were susceptible to the antibiotic and could not regenerate. To check if this was the case,

kill curves with lower concentrations of spectinomycin dihydrochloride (50, 100, 200, 300, 400 and 500mg/ml) can be used with each of the Nt6495-30, Nt6495-16 and Nt6495-69 lines to establish the concentration at which the regeneration of shoots are as good as in wild type. This concentration of spectinomycin dihydrochloride will then be used to repeat transformation frequency tests with the three Nt6495 lines.

To analyze for direct glyphosate selection, kill curves with varying levels of glyphosate will be established with the Nt6495 lines to find the best selection level. Plastid transformation vector pMON49218 will be used to bombard the Nt6495 lines and tested for direct selection using the optimized glyphosate level.

In Arabidopsis, FtsZ1 nuclear expression construct pCGN6495 was used to transform Columbia ecotype. T1 seeds were collected and about 100 kanamycin resistant seedlings were analyzed for alteration in plastid size and number following the same protocols as outlined for the tobacco section of this report. The transgenic plants were divided into three groups based on plastid number—I) 20 independent lines containing few large plastid (1-5),(II) 23 lines lines containing 5-20 plastids and (III) 50 lines containing wild-type plastid number were obtained. Selected T2 plants from each category were analyzed for number of transgene integration loci and sent to the growth chamber for T3 seed collection to identify homozygous plants. Such plants can be used in plastid transformations as described by Sikdar, *et al.* (1998) *Plant Cell Reports*, 18:20-24.

Transformed plants selected for *aad*A marker gene expression or glyphosate resistance are analyzed to determine whether the entire plastid content of the plant has been transformed (homoplasmic transformants). Typically, following two rounds of shoot formation and spectinomycin selection, approximately 50% of the transgenic plantlets which are analyzed are homoplasmic, as determined by Southern blot analysis of plastid DNA. Homoplasmic plantlets are selected for further cultivation.

Southern blot analysis is used to confirm the integration of the chimeric expression cassettes in the plastid genome. Preparation, electrophoresis, and transfer of DNA to filters is as described (Svab *et al.*, (1993 *supra*)). Total plant cellular DNA can be prepared as described by Dellaporta *et al.* (1983) *Plant Mol. Biol. Rep. 1*:19-21).

To visually observe the expression of marker genes such as GFP from the chloroplasts of transformed plants, various tissues are visualized utilizing a dissecting microscope. Protoplasts and chloroplasts are isolated as described in Sidorov, *et al.* (1994) *Theor. Appl. Genet.* 88:525-529.

The above results demonstrate that the sequences of the present invention provide an efficient means for the production of plastid transformed plants. Furthermore, such methods find use in plastid transformation methods involving the selection of transplastomic plants on herbicides, for example glyphosate.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claim.

CLAIMS

What is claimed is:

1. In a method for transforming a plant cell plastid comprising the steps of introducing into cells of a plant a construct comprising a promoter functional in a plant cell plastid operably associated with a DNA sequence of interest and transforming said plant cell plastid with said construct,

wherein the improvement comprises introducing said construct into a plant cell having an altered plant plastid morphology selected from the group consisting of altered plastid size and altered plastid number in said plant cell.

2. The method according to Claim 1, wherein said plastid size is increased from a wild-type plant plastid morphology.

3. The method according to Claim 2, further wherein said plastid number is decreased from a wild-type plant plastid morphology.

4. The method according to Claim 1, wherein said plastid size is decreased from a wild-type plant plastid morphology.

5. The method according to Claim 4, further wherein said plastid number is increased from a wild-type plant plastid morphology.

6. The method according to Claim 2, wherein said plant cell is obtained from a plant tissue source in which plastid division is inhibited.

7. The method according to Claim 6, wherein said plastid division is inhibited by introduction into cells of the plant tissue source a second DNA construct comprising in the 5' to 3' direction of transcription a promoter functional in a plant cell, a DNA sequence coding for a gene involved in plastid cell division and a transcriptional termination sequence functional in a plant cell.

8. The method according to Claim 7, wherein said DNA sequence is in an antisense orientation.

9. The method according to Claim 8, wherein said construct contains a DNA sequence coding for an FtsZ protein.

10. The method according to Claim 7, wherein said DNA sequence is in a sense orientation.

11. The method according to Claim 10, wherein said DNA sequence provides for sense suppression.

12. The method according to Claim 6, wherein said plastid division is inhibited by growing a plant under culture conditions which inhibit the division of plant cell plastids.

13. The method according to Claim 12, wherein said culture conditions comprise growing the plant tissue source under exposure to an inhibitor of bacterial cell division.

14. The method according to Claim 13, wherein said inhibitor is 5,5'-Bis-(8-anilino-1-naphthalenesulfonate).

15. The method according to Claim 6, wherein said plastid division is inhibited by genetic mutagenesis.

16. An isolated DNA sequence encoding a plant FtsZ protein from *Arabidopsis thaliana*.

17. The DNA sequence of Claim 16, wherein said FtsZ protein is encoded by a sequence which includes a sequence selected from the group consisting of SEQ ID Nos:1 and 3.

18. An isolated DNA sequence encoding a plant FtsZ protein from *Brassica* .

19. The DNA sequence of Claim 18, wherein said FtsZ protein is encoded by a sequence of SEQ ID NO:5.

20. An isolated DNA sequence encoding a plant FtsZ protein from soybean.

21. The DNA sequence of Claim 20, wherein said FtsZ protein is encoded by a sequence which includes a seqeunce selected from the group consisting of SEQ ID NOs:20-31.

22. An isolated DNA sequence encoding a plant FtsZ protein from corn.

23. The DNA encoding sequence of Claim 22, wherein said FtsZ protein is encoded by a sequence which includes a sequence selected from the group consisting of SEQ ID Nos:10-19.

24. A recombinant DNA construct comprising any of the DNA encoding sequences of Claims 16-23.

25. A plant cell comprising the DNA construct of 24.

26. A plant comprising a cell of Claim 25.

27. A method for improving the selectability of plant comprising.

transforming a plant cell source having an altered plastid morphology with a construct comprising a promoter functional in a plant cell plastid operably associated with a nucleic acid sequence encoding a selectable marker.

28. The method according to Claim 27, wherein said nucleic acid sequence encodes an herbicide tolerance gene.

29. The method according to Claim 27, wherein said nucleic acid sequence encodes a glyphosate tolerance gene.

30. A method for preparing a plant cell source with increased plastid transformation efficiency comprising,

transforming a plant cell with a construct comprising a promoter functional in plant cell operably associated with a nucleic acid sequence encoding a FtsZ protein.

31. A method for transforming a plant cell plastid comprising,

introducing into a plant cell having altered plastid morphology a first nucleic acid construct comprising a promoter functional in a plant cell plastid operably associated with a nucleic acid sequence of interest.

ClustalW Formatted Alignments

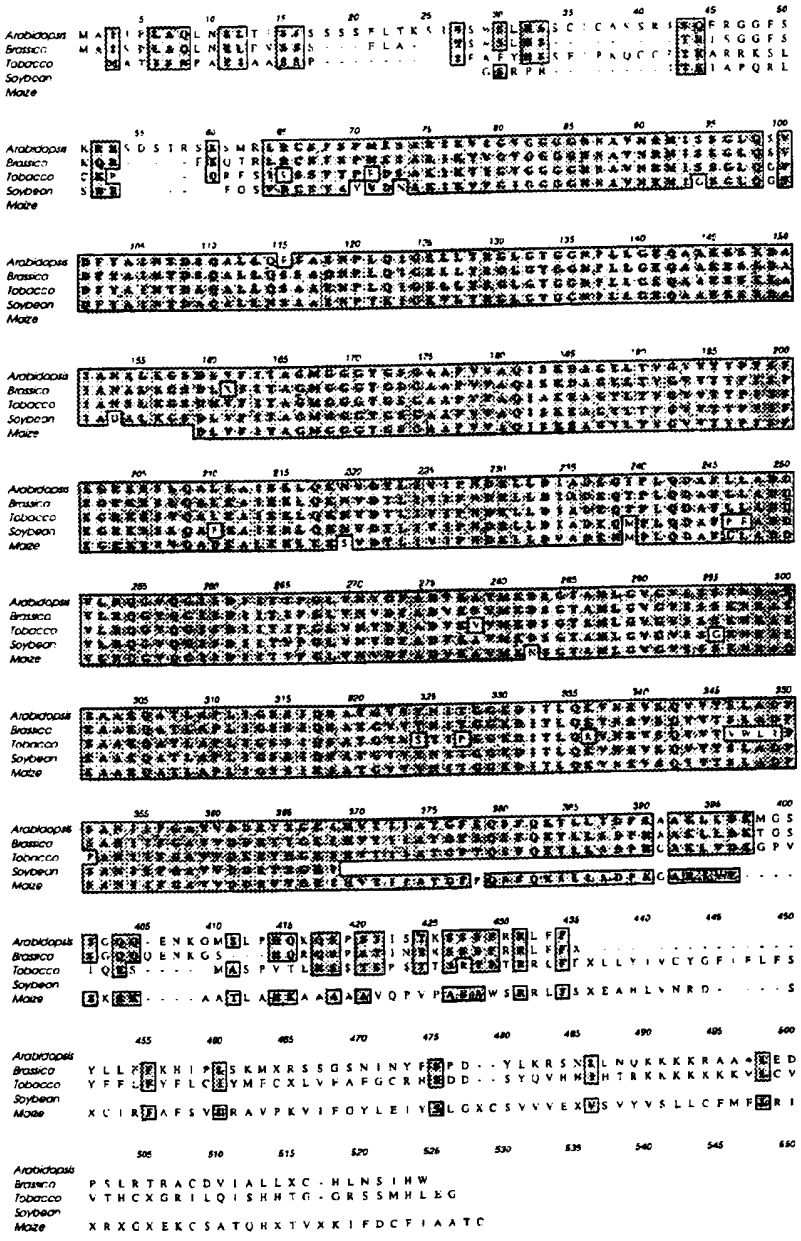

Figure 1

SEQUENCE LISTING

<110> Calgene LLC

<120> Methods for Transforming Plastids

<130> 15595/00/WO

<150> 60/109,892
<151> 1998-11-25

<160> 73

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1495
<212> DNA
<213> Arabidopsis sp

<400> 1
atggcgataa ttccgttagc acagcttaat gagctaacga tttcttcatc ttcttcttcg 60
tttcttacca aatcgatatc ttctcattcg ttgcatagta gctgcatttg cgcaagttct 120
agaatcagtc aattccgtgg cggcttctct aaacgaagaa gcgattcaac aaggtctaag 180
tcgatgcgat tgaggtgttc cttctctccg atggaatctg cgagaattaa ggtgattggt 240
gtcggtggtg gtggtaacaa tgccgttaac cggatgattt caagcggttt acagagtgtt 300
gatttctatg cgataaacac ggattcgcaa gctctgttac agtcttctgc tgagaaccca 360
cttcaaattg gagaactttt aactcgtggg cttggcactg gtggaaaccc gcttcttgga 420
gaacaagctg cagaagaatc aaaagatgca attgctaatg ctcttaaagg atcagacctt 480
gttttcataa ctgctggtat gggtggtgga acagggtctg gtgctgcacc tgtggtagct 540
cagatttcga aggatgctgg ttatttgact gttggtgttg ttacctatcc gtttagcttt 600
gaaggacgta aaagatcttt gcaggcactg gaagctattg aaaagctcca aaagaatgtt 660
gataccctta tcgtgattcc aaatgatcgt ctgctagata ttgctgatga acagacgcca 720
cttcaggacg cgtttcttct tgcagatgat gttttacgcc aaggagtaca aggaatctca 780
gatattatta ctatacctgg actagtcaat gtagattttg cagatgtgaa ggcagtcatg 840
aaagattctg gaactgcaat gctcggggta ggtgtttctt ccagcaaaaa ccgggcagaa 900
gaagcagctg aacaagcaac tttggctcca ttgatcggat catccataca atcagctact 960
ggtgtcgtct acaacatcac tggtggaaaa gacataactt tgcaggaagt gaaccgagta 1020
tcacaggtcg tgacaagttt ggcagaccca tcggccaaca tcatatttgg agctgttgtg 1080
gatgatcgct acaccggaga gattcatgta acgataatcg ccacaggctt ctctcagtca 1140
ttccagaaga cacttctgac tgatccaaga gcagctaaac tccttgacaa aatgggatca 1200
tcaggtcaac aagagaacaa aggaatgtct ctgcctcacc agaagcagtc tccatcaact 1260
atctctacca aatcgtcttc tccccgtaga cttttcttct agttttcttt ttttcctttt 1320
cggtttcaag catcaaaaat gtaacgatct tcaggctcaa atatcaatta catttgattt 1380
tcaaaaaaaa aaaaaaaagg cggccgctct agaggatcca agcttacgta cgcgtgcatg 1440
cgacgtcata gctcttctat agtgtcacct aaattcaatt cactggccgt cgttt 1495

<210> 2
<211> 433
<212> PRT
<213> Arabidopsis sp

<400> 2

Met Ala Ile Ile Pro Leu Ala Gln Leu Asn Glu Leu Thr Ile Ser Ser
1 5 10 15
Ser Ser Ser Ser Phe Leu Thr Lys Ser Ile Ser Ser His Ser Leu His
20 25 30
Ser Ser Cys Ile Cys Ala Ser Ser Arg Ile Ser Gln Phe Arg Gly Gly
35 40 45
Phe Ser Lys Arg Arg Ser Asp Ser Thr Arg Ser Lys Ser Met Arg Leu
50 55 60
Arg Cys Ser Phe Ser Pro Met Glu Ser Ala Arg Ile Lys Val Ile Gly
65 70 75 80
Val Gly Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Ser Ser Gly
85 90 95
Leu Gln Ser Val Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu
100 105 110
Leu Gln Phe Ser Ala Glu Asn Pro Leu Gln Ile Gly Glu Leu Leu Thr
115 120 125
Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala
130 135 140
Glu Glu Ser Lys Asp Ala Ile Ala Asn Ala Leu Lys Gly Ser Asp Leu
145 150 155 160
Val Phe Ile Thr Ala Gly Met Gly Gly Gly Thr Gly Ser Gly Ala Ala
165 170 175
Pro Val Val Ala Gln Ile Ser Lys Asp Ala Gly Tyr Leu Thr Val Gly
180 185 190
Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Leu Gln
195 200 205
Ala Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile
210 215 220
Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Gln Thr Pro
225 230 235 240
Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val
245 250 255
Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp
260 265 270
Phe Ala Asp Val Lys Ala Val Met Lys Asp Ser Gly Thr Ala Met Leu
275 280 285
Gly Val Gly Val Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu
290 295 300
Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr
305 310 315 320
Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu
325 330 335
Val Asn Arg Val Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala
340 345 350
Asn Ile Ile Phe Gly Ala Val Val Asp Asp Arg Tyr Thr Gly Glu Ile
355 360 365
His Val Thr Ile Ile Ala Thr Gly Phe Ser Gln Ser Phe Gln Lys Thr
370 375 380
Leu Leu Thr Asp Pro Arg Ala Ala Lys Leu Leu Asp Lys Met Gly Ser
385 390 395 400
Ser Gly Gln Gln Glu Asn Lys Gly Met Ser Leu Pro His Gln Lys Gln
405 410 415
Ser Pro Ser Thr Ile Ser Thr Lys Ser Ser Ser Pro Arg Arg Leu Phe
420 425 430
Phe

<210> 3
<211> 1611
<212> DNA
<213> Arabidopsis sp

<400> 3
tgttgttgcc gctcagaaat ctgaatcttc tccaatcaga aactctccac ggcattacca 60
aagccaagct caagatcctt tcttgaacct tcacccggaa atatctatgc ttagaggtga 120
agggactagt acaatagtca atccaagaaa ggaaacgtct tctggaccta ttgtcgagga 180
ttttgaagag ccatctgctc cgagtaacta caatgaggcg aggattaagg ttattggtgt 240
gggaggtggt ggatcaaatg ctgtgaatcg tatgatagag agtgaaatgt caggtgtgga 300
gttctggatt gtcaacactg atatccaggc tatgagaatg tctcctgttt tgcctgataa 360
taggttacaa attggtaagg agttgactag gggtttaggt gctggaggaa atccagaaat 420
cggtatgaat gctgctagag agagcaaaga agttattgaa gaagctcttt atggctcaga 480
tatggtcttt gtcacagctg gaatgggcgg tggaactggc actggtgcag cccctgtaat 540
tgcaggaatt gccaaggcga tgggtatatt gacagttggt attgccacaa cgcctttctc 600
gtttgagggt cgaagaagaa ctgttcaggc tcaagaaggg cttgcatctc tcagagacaa 660
tgttgacact ctcatcgtca ttccaaatga caagttgctt acagctgtct ctcagtctac 720
tccggtaaca gaagcattta atctagctga tgatatactc cgtcaggggg ttcgtgggat 780
atctgatatc attacgattc ctggtttggt gaatgtggat tttgctgatg tgagagctat 840
aatggcaaat gcggggtctt cattgatggg aataggaact gcgacaggaa agagtcgggc 900
aagagatgct gcgctaaatg caatccaatc ccctttgtta gatattggga ttgagagagc 960
cactggaatt gtttggaaca ttactggcgg aagtgacttg acattgtttg aggtaaatgc 1020
tgctgcggaa gtaatatatg atcttgtcga tccaactgcc aatcttatat tcggtgctgt 1080
tgtagatcca gccctcagcg gtcaagtaag cataaccctg atagctacgg gtttcaaacg 1140
acaagaagag ggagaaggac gaacagttca gatggtacaa gcagatgctg cgtcagttgg 1200
agctacaaga agaccctctt cttcctttag agaaagcggt tcagtggaga tcccagagtt 1260
cttgaagaag aaaggcagct ctcgttatcc ccgagtctaa agcccaatct aatcactacc 1320
ctgcacactg cagcaataac aaacgtgtgt gtactggtag tctggtactg ccttctggga 1380
tacagcaaga tgtgttgatg tatgatcaag aatctgtgtg ggtgtgtata tgttctgtca 1440
ctgcctctgg tcgtgttctt gaataggttg ttttagaaat cggagtttct ctctatgtca 1500
cttccaaaac aaaaaaggag aagaagaatc acacttctcg aaccataaac atacttataa 1560
gattatgaga gttttagcag aaattattgt caaaaaaaaa aaaaaaaaaa a 1611

<210> 4
<211> 397
<212> PRT
<213> Arabidopsis sp

<400> 4
Met Leu Arg Gly Glu Gly Thr Ser Thr Ile Val Asn Pro Arg Lys Glu
1 5 10 15
Thr Ser Ser Gly Pro Val Val Glu Asp Phe Glu Glu Pro Ser Ala Pro
20 25 30
Ser Asn Tyr Asn Glu Ala Arg Ile Lys Val Ile Gly Val Gly Gly Gly
35 40 45
Gly Ser Asn Ala Val Asn Arg Met Ile Glu Ser Glu Met Ser Gly Val
50 55 60
Glu Phe Trp Ile Val Asn Thr Asp Ile Gln Ala Met Arg Met Ser Pro
65 70 75 80
Val Leu Pro Asp Asn Arg Leu Gln Ile Gly Lys Glu Leu Thr Arg Gly
85 90 95
Leu Gly Ala Gly Gly Asn Pro Glu Ile Gly Met Asn Ala Ala Arg Glu

```
            100                 105                 110
Ser Lys Glu Val Ile Glu Glu Ala Leu Tyr Gly Ser Asp Met Val Phe
        115                 120                 125
Val Thr Ala Gly Met Gly Gly Gly Thr Gly Thr Gly Ala Ala Pro Val
    130                 135                 140
Ile Ala Gly Ile Ala Lys Ala Met Gly Ile Leu Thr Val Gly Ile Ala
145                 150                 155                 160
Thr Thr Pro Phe Ser Phe Glu Gly Arg Arg Arg Thr Val Gln Ala Gln
                165                 170                 175
Glu Gly Leu Ala Ser Leu Arg Asp Asn Val Asp Thr Leu Ile Val Ile
            180                 185                 190
Pro Asn Asp Lys Leu Leu Thr Ala Val Ser Gln Ser Thr Pro Val Thr
        195                 200                 205
Glu Ala Phe Asn Leu Ala Asp Asp Ile Leu Arg Gln Gly Val Arg Gly
    210                 215                 220
Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp Phe Ala
225                 230                 235                 240
Asp Val Arg Ala Ile Met Ala Asn Ala Gly Ser Ser Leu Met Gly Ile
                245                 250                 255
Gly Thr Ala Thr Gly Lys Ser Arg Ala Arg Asp Ala Ala Leu Asn Ala
            260                 265                 270
Ile Gln Ser Pro Leu Leu Asp Ile Gly Ile Glu Arg Ala Thr Gly Ile
        275                 280                 285
Val Trp Asn Ile Thr Gly Gly Ser Asp Leu Thr Leu Phe Glu Val Asn
    290                 295                 300
Ala Ala Ala Glu Val Ile Tyr Asp Leu Val Asp Pro Thr Ala Asn Leu
305                 310                 315                 320
Ile Phe Gly Ala Val Val Asp Pro Ala Leu Ser Gly Gln Val Ser Ile
                325                 330                 335
Thr Leu Ile Ala Thr Gly Phe Lys Arg Gln Glu Glu Gly Glu Gly Arg
            340                 345                 350
Thr Val Gln Met Val Gln Ala Asp Ala Ala Ser Val Gly Ala Thr Arg
        355                 360                 365
Arg Pro Ser Ser Ser Phe Arg Glu Ser Gly Ser Val Glu Ile Pro Glu
    370                 375                 380
Phe Leu Lys Lys Lys Gly Ser Ser Arg Tyr Pro Arg Val
385                 390                 395
```

<210> 5
<211> 1450
<212> DNA
<213> Brassica sp

<400> 5

```
atggcgatta gtccgttggc acagcttaac gagctaccag tctcttcctc gtttcttgcg      60
acatcccact cgctgcacag taccagaatc agtggcggct tctcaaaaca aaggtttaag     120
caaacacggt tgagatgctc cttctctccg atggagtctg cgaggattaa ggtggttggt     180
gtcggcggtg gtggtaacaa tgccgtcaat cgcatgattt ccagcggctt acagagtgtt     240
gatttctatg cgataaacac ggactctcaa gctctcttgc agtcttctgc gcagaaccct     300
cttcaaattg gagagctcct aactcgtggc cttgggactg gtgggaaccc gcttctagga     360
gaacaagctg ctgaggaatc taaagacgcg attgctaatg ctcttaaagg atctgacctt     420
gytttcatta ctgctggtat gggtggtggc actggctccg gtgctgctcc tgttgttgct     480
cagatctcga aagacgctgg ttatttgacc gttggtgttg ttacctatcc cttcagcttc     540
gaaggtcgta aaagatcttt gcaggcactt gaagccattg aaaagctgca gaagaacgtg     600
gataccctca tcgtgatacc aaatgatcgt ctcctagata ttgctgatga acagacgcct     660
```

```
cttcaagacg cttttcttct cgccgatgat gttttgcggc aaggagttca aggaatctct     720
gatattatta ctatacctgg actggtcaat gtagattttg cggatgtgaa gtcggttatg     780
aaagattccg gaactgcgat gctcggggtg ggtgtttctt caagcaagaa ccgagcagaa     840
gaagcagctg agcaagccac tttggctcca ttgatcggat catccattca atcagctact     900
ggtgtcgtct acaacatcac cggtggaaaa gacattactt tgcaggaagt gaaccgagta     960
tctcaggtgg tgacaagttt ggcagaccca tcggccaaca tcatatttgg agctgttgtg    1020
gatgatcgat acactggaga gattcatgta acgataatag ccacggggtt ctcacagtct    1080
ttccagaaga cacttctcag tgatccaaga gcagctaaac tactcgacaa aacgggatca    1140
tcaggtcaac aacaagagaa caaaggcagt caccagaggc agtctcctgc aactatcaac    1200
accaaatcat cttctccccg tagattgttc ttctagtatc ttttgttttt taagcatatt    1260
cctttatcaa aaatgtaacg atcttcaggc tcaaatatca attacttttc tccagattat    1320
ctcaaaagaa gtaatttgtt aaaccaaaaa aaaaaaaaaa gggcggccgc tctagaggat    1380
ccaagcttac gtacgcgtgc atgcgacgtc atagctcttc tatagtgtca cctaaattca    1440
attcactggc                                                          1450
```

<210> 6
<211> 411
<212> PRT
<213> Brassica sp

<220>
<221> VARIANT
<222> (1)...(411)
<223> Xaa = Any Amino Acid

<400> 6

```
Met Ala Ile Ser Pro Leu Ala Gln Leu Asn Glu Leu Pro Val Ser Ser
 1               5                  10                  15
Ser Phe Leu Ala Thr Ser His Ser Leu His Ser Thr Arg Ile Ser Gly
            20                  25                  30
Gly Phe Ser Lys Gln Arg Phe Lys Gln Thr Arg Leu Arg Cys Ser Phe
        35                  40                  45
Ser Pro Met Glu Ser Ala Arg Ile Lys Val Val Gly Val Gly Gly Gly
    50                  55                  60
Gly Asn Asn Ala Val Asn Arg Met Ile Ser Ser Gly Leu Gln Ser Val
65                  70                  75                  80
Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu Leu Gln Ser Ser
                85                  90                  95
Ala Gln Asn Pro Leu Gln Ile Gly Glu Leu Leu Thr Arg Gly Leu Gly
            100                 105                 110
Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala Glu Glu Ser Lys
        115                 120                 125
Asp Ala Ile Ala Asn Ala Leu Lys Gly Ser Asp Leu Xaa Phe Ile Thr
    130                 135                 140
Ala Gly Met Gly Gly Gly Thr Gly Ser Gly Ala Ala Pro Val Val Ala
145                 150                 155                 160
Gln Ile Ser Lys Asp Ala Gly Tyr Leu Thr Val Gly Val Val Thr Tyr
                165                 170                 175
Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Leu Gln Ala Leu Glu Ala
            180                 185                 190
Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile Val Ile Pro Asn
        195                 200                 205
Asp Arg Leu Leu Asp Ile Ala Asp Glu Gln Thr Pro Leu Gln Asp Ala
    210                 215                 220
Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val Gln Gly Ile Ser
```

```
225                 230                 235                 240
Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp Phe Ala Asp Val
                245                 250                 255
Lys Ser Val Met Lys Asp Ser Gly Thr Ala Met Leu Gly Val Gly Val
            260                 265                 270
Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu Gln Ala Thr Leu
        275                 280                 285
Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr Gly Val Val Tyr
    290                 295                 300
Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu Val Asn Arg Val
305                 310                 315                 320
Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala Asn Ile Ile Phe
                325                 330                 335
Gly Ala Val Val Asp Asp Arg Tyr Thr Gly Glu Ile His Val Thr Ile
            340                 345                 350
Ile Ala Thr Gly Phe Ser Gln Ser Phe Gln Lys Thr Leu Leu Ser Asp
        355                 360                 365
Pro Arg Ala Ala Lys Leu Leu Asp Lys Thr Gly Ser Ser Gly Gln Gln
    370                 375                 380
Gln Glu Asn Lys Gly Ser His Gln Arg Gln Ser Pro Ala Thr Ile Asn
385                 390                 395                 400
Thr Lys Ser Ser Ser Pro Arg Arg Leu Phe Phe
                405                 410
```

<210> 7
<211> 1295
<212> DNA
<213> Nicotiana sp

<400> 7

```
tgccgttaac cggatgattt caagcggttt acagggtgtt gacttctatg ctataaacac      60
ggatgctcaa gcactgctac agtctgctgc tgaaaacccg cttcaaattg gagaacttct     120
gactcgtggg cttggtactg gtggtaatcc tctttttaggg gaacaggcag tggaggagtc     180
gaaggaagcc attgcaaatt ctctaaaagg ttcagatatg gtgttcataa cagcaggaat     240
gggtggaggt acaggatctg gtgctgctcc tgttgtggct caaatagcaa aagaagcagg     300
ctatttgact gttggtgttg tcacataccc attcagcttt gaaggacgta aaagatccgt     360
gcaggctctg gaagcaattg aaaaacttca gaaaaatgta gataccctta tagtaattcc     420
aatgaccgtc tgctagatat tgctgatgag cagacaccac ttcaagatgc ttttcttctt     480
gctgatgatg tattacgcca aggtgtccaa ggaatttccg atataattac tatacctggg     540
cttgtaaatg tggattttgc cgatgtaaag gtagtgatga aagattctgg aactgctatg     600
cttggagttg gggtttcatc aagcaagaac cgtgctgaag aagcagccga acaagcaact     660
cttgcccctc ttaattggat cgtccattca atcgccactg gggtagtatc caccattcca     720
ggaggaaaag accataactt tgcagaaagt gaatagggtg tctcaggttg ttacagtctg     780
gctgatccct cccgctaaca tcatatttgg tgctgttgtg gatgagcgct acaatggcga     840
aatacacgtg accataattg caactggttt tacccagtct tttcagaaga ctcttctctc     900
tgacccacga ggtgcaaagc ttgttgataa aggcccagta atccaagaaa gcatggcatc     960
acctgttacc ctgaggtcat caacctcacc ttcgacaaca tcacgaacac ctactcggag    1020
gctgttcttt tagctccttt atatagtttg ttacggcttc atttttctct tttcttactt    1080
ttttcttttt tactttcttt gtatttacat gttttgctga ttggtgtttg catttggctg    1140
tagacatagt gatgattctt atcaagtgca tcacattcat actcgaaaaa aaaaaaaaaa    1200
aaaaaaagta ctctgcgttg ttacccactg ttaagggcga attctgcaga tatcccatca    1260
cactggcggc cgctcgagca tgcatctaga gggcc                               1295
```

<210> 8
<211> 1255

<212> DNA
<213> Nicotiana sp

<400> 8
atggccacca tctcaaaccc agcagagata gcagcttctt ctccttcctt tgctttttac 60
cactcttcct ttattcctaa acaatgctgc ttcaccaaag ctcgccggaa aagcttatgt 120
aaacctcaac gtttcagcat ttcaagttca tttactcctt ttgattctgc taagattaag 180
gttatcggcg tcggtggcgg tggtaacaat gccgttaacc ggatgatttc aagcggttta 240
cagggtgttg acttctatgc tataaacacg gatgctcaag cactgctgca gtctgctgct 300
gaaaacccgc ttcaaattgg agaacttctg actcgtgggc ttggtactgg tggtaatcct 360
cttttagggg aacaggcagc ggaggagtcg aaggaagcca ttgcaaattc tctaaaaggt 420
tcagatatgg tgttcataac agcaggaatg ggtggaggta caggatctgg tgctgctcct 480
gttgtggctc aaatagcaaa agaagcaggc tatttgactg ttggtgttgt cacataccca 540
ttcagctttg aaggacgtaa aagatccgtg caggctctgg aagcaattga aaaacttcag 600
aaaaatgtag atacccttat agtaattccc aatgaccgtc tgctagatat tgctgatgag 660
cagacaccac ttcaagatgc ttttcttctt gctgatgatg tattacgcca aggtgtccaa 720
ggaatttccg atataattac tatacctggg cttgtaaatg tggattttgc cgatgtaaag 780
gtagtgatga aagattctgg aactgctatg cttggagttg gggtttcatc aagcaagaac 840
cgtgctgaag aagcagccga acaagcaact cttgcccctc ttattggatc gtccattcaa 900
tcagccactg gggtagtatc caccattcca ggaggaaaag acataacttt gcagaaagtg 960
aatagggtgt ctcaggttgt tacagtctgg ctgatccctc ccgctaacat catatttggt 1020
gctgttgtgg atgagcgcta caatggcgaa atacacgtga ccataattgc aactggtttt 1080
acccagtctt ttcagaagac tcttctctct gacccacgag gtgcaaagct tgttgataaa 1140
ggcccagtaa tccaagaaag catggcatca cctgttaccc tgaggtcatc aacctcacct 1200
tcgacaacat cacgaacacc tactcggagg ctgttctttt agctccttta tatag 1255

<210> 9
<211> 413
<212> PRT
<213> Nicotiana sp

<400> 9
Met Ala Thr Ile Ser Asn Pro Ala Glu Ile Ala Ala Ser Ser Pro Ser
1 5 10 15
Phe Ala Phe Tyr His Ser Ser Phe Ile Pro Lys Gln Cys Cys Phe Thr
20 25 30
Lys Ala Arg Arg Lys Ser Leu Cys Lys Pro Gln Arg Phe Ser Ile Ser
35 40 45
Ser Ser Phe Thr Pro Phe Asp Ser Ala Lys Ile Lys Val Ile Gly Val
50 55 60
Gly Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Ser Ser Gly Leu
65 70 75 80
Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr Asp Ala Gln Ala Leu Leu
85 90 95
Gln Ser Ala Ala Glu Asn Pro Leu Gln Ile Gly Glu Leu Leu Thr Arg
100 105 110
Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala Glu
115 120 125
Glu Ser Lys Glu Ala Ile Ala Asn Ser Leu Lys Gly Ser Asp Met Val
130 135 140
Phe Ile Thr Ala Gly Met Gly Gly Gly Thr Gly Ser Gly Ala Ala Pro
145 150 155 160
Val Val Ala Gln Ile Ala Lys Glu Ala Gly Tyr Leu Thr Val Gly Val
165 170 175
Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Val Gln Ala

```
            180                 185                 190
Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile Val
        195                 200                 205
Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Gln Thr Pro Leu
    210                 215                 220
Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val Gln
225                 230                 235                 240
Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp Phe
                245                 250                 255
Ala Asp Val Lys Val Val Met Lys Asp Ser Gly Thr Ala Met Leu Gly
            260                 265                 270
Val Gly Val Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu Gln
        275                 280                 285
Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr Gly
    290                 295                 300
Val Val Ser Thr Ile Pro Gly Gly Lys Asp Ile Thr Leu Gln Lys Val
305                 310                 315                 320
Asn Arg Val Ser Gln Val Val Thr Val Trp Leu Ile Pro Pro Ala Asn
                325                 330                 335
Ile Ile Phe Gly Ala Val Val Asp Glu Arg Tyr Asn Gly Glu Ile His
            340                 345                 350
Val Thr Ile Ile Ala Thr Gly Phe Thr Gln Ser Phe Gln Lys Thr Leu
        355                 360                 365
Leu Ser Asp Pro Arg Gly Ala Lys Leu Val Asp Lys Gly Pro Val Ile
    370                 375                 380
Gln Glu Ser Met Ala Ser Pro Val Thr Leu Arg Ser Ser Thr Ser Pro
385                 390                 395                 400
Ser Thr Thr Ser Arg Thr Pro Thr Arg Arg Leu Phe Phe
                405                 410
```

<210> 10
<211> 1278
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(1278)
<223> n = A,T,C or G

<400> 10

```
gatcttgtct tcataacagc tgggatggga gggggtactg gatctggtgc tgctccagtt      60
gttgcccaga tatcaaagga agctggttat cttactgttg gtgttgtcac ctatccattc     120
agtttcgagg gccgtaagcg ctctgtacag gcattggaag cactagagaa gctggaaaag     180
agtgtagaca cacttattgt gattccaaat gataagttat tagatgttgc ggatgaaaac     240
atgcccttgc aagatgcatt tctccttgca gatgatgtcc ttcgtcaggg tgttcaagga     300
atatcagaca tcatcacaat accgggactt gtcaatgttg attttgctga tgtaaaagct     360
gtcatgaaaa actctggaac tgccatgctc ggtgttggtg tttcttccag caaaaatcgg     420
gcccaagaag ctgctgaaca ggcaacactt gctcctttga ttggatcatc catcgaggca     480
gctactggcg ttgtgtataa tattactggt gggaaggaca tcactttgca agaagtgaac     540
aaggtgtccc agattgtgac aagcctagct gacccatctg cgaacataat ttcggtgct      600
gtcgttgatg accgttacac tggtgagata catgtgacaa tcattgcgac aggatttcca     660
cagtccttcc agaaatccct tttggcggat ccaaagggag cacgtatagt ggaatccaaa     720
gagaaagcag caaccctcgc ccataaagca gcagcagctg cagttcaacc ggtccctgct     780
tctgcttggt ctcgaagact cttctcctga gaagctcatt tggtgaaccg tgactcgtag     840
```

tgcattagat ttgcatttag cgtgttgagg gcagtcccta aggtgatctt cggatatctg 900
gagatttata gcttgggcta gtgttcggta gtggtagaat aagtttcagt gtatgtatcg 960
ttgctttgct ttatgttttt gaggatcagg cggtgaggct gagagaagtg ctcagcaact 1020
caacattgaa ctgttgtaga agatctttga ttgcttttat tgctgcaaca tgccaacaac 1080
cctctgttgg attcamcmna aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1140
aaaaaaaaaa aaaaaaaaaa aaaannncaaa aaaaaaaaaa aaaaaaaagg gcggccgccg 1200
actagtgagc tcgtcgaccc gggaattaat tccggaccgg tacctgcagg cgtaccagct 1260
ttccctatag tgagtcgt 1278

<210> 11
<211> 283
<212> DNA
<213> Zea mays

<400> 11
gctccagttg ttgcccagat atcaaaggaa gctggttatc ttactgttgg tgttgtcacc 60
tatccattca gtttcgaggg ccgtaagcgc tctgtacagg cattggaagc actagagaag 120
ctggaaaaga gtgtagacac acttattgtg attccaaatg ataagttatt agatgttgcg 180
gatgaaaaca tgcccttgca agatgcattt ctccttgcag atgatgtcct tcgtcagggt 240
gttcaaggaa tatcagacat catcacaata ccgggacttg tca 283

<210> 12
<211> 287
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(287)
<223> n = A,T,C or G

<400> 12
gggccgtaag cgctctgtac aggcattgga agcactagag aagctggaaa agagtgtaga 60
cacacttatt gtgattccan atnatnngtt attagatgtt gcggatgaaa acatgccctt 120
gcaagatgca tttctccttg cagatgatgt ccttcgtcag ggtgttcaag gaatatcaga 180
catcatcaca ataccgggac ttgtcaatgt tgattttgct gatgtaaaag ctgtcatgaa 240
aaactctgga actgccatgc tcggtgttgg tgtttcttcc agcaaaa 287

<210> 13
<211> 1122
<212> DNA
<213> Zea mays

<400> 13
gctataaaca ccgattccca agcccttatt aattcacaag cgcaatatcc tctgcaaatt 60
ggagagcagt tgacccgcgg cttaggtgcc ggtggaaatc cgaatttggg agagcaggct 120
gctgaggaat caagagaaac catagccact gccctgaggg attcagatct tgtcttcata 180
acagctggga tgggaggggg tactggatct ggtgctgctc cagttgttgc ccagatatca 240
aaggaagctg gttatcttac tgttggtgtt gtcacctatc cattcagttt cgagggccgt 300
aagcgctctg tacaggcaaa gtatctgagc ccccccttcac tcctgaattt taattcaaac 360
tgtcatatct cgttctgcga ctttcttttg ctcgatggaa gcattagttt gtagtcataa 420
caatgacatc cagccacatt tattgctgat gatgtataca atggtaggtc aaagaaatgt 480
agcatcatgc catcacctgt agttcatctc atcattttgt tcctactttt ctgcgtggtt 540
gatgcccaaa acaatataca actatgtggt tgtactgttg cattgccttg tggagggatg 600
tttatgttgt gaaatatttc aaaacacatg tcattatgaa tattccctcc tgtggttgtg 660

gggacttgtt tcaaatgcta tgaattaaga acaaggcaac ataaagtgtt aaatgttaac 720
cgtctttcgt ccatgaaaca ttattccctt gaggataatg ggccttggac aaaggctgat 780
gagagtataa ttaccaagct taaatcttcg taataaaatt tcaatagata ttgtaagata 840
acataaaata aagggtataa aaaggggtaa ataaatcata gacgaattat attatattta 900
cttaatatat tgaatcattg aatacaataa tacctctgcc ttggcaaagg ttggattccg 960
aaaaatgtga ttgcaagtta ccagaatgcg tgaacagtaa aggaatactg ttcactattt 1020
ataggcacag gacacagcct gtggaggaat tcaattatac ccgtcataag agtttacaca 1080
ttgacttaga cctttatgga ctaaaagatc attgctatct tt 1122

<210> 14
<211> 291
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(291)
<223> n = A,T,C or G

<400> 14
aaaatagtgt ggacacccta atcgtcatcc caaatgataa gttgctgtct gctgtttctc 60
caaatacacc tgtaactgaa gcatttaatc tggctgatga tattcttcgt caaggcattc 120
gtggcatatc tgatataatt acggttcctg ggnaggttaa tgttgatttt gctgacgtac 180
gtgctatcat gcaaaatgca gggtcatcct tgatgggtat agggactgct acaggaaagt 240
caagagcaag ggatgctgct cttaacgcca tccagtcgcc gctgcttgat a 291

<210> 15
<211> 415
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(415)
<223> n = A,T,C or G

<400> 15
gagcaaggga tgctgctctt aacgccatcc agtcgccgct gcttgatatt ggaattgaaa 60
gagccacagg cattgtgtgg aatatcactg ggggaactga cctgactttg tttgaggtga 120
atgctgcggc cgaaattatc tacgaccttg tcgatccaaa cgctaatctg atatttggcg 180
ccgtcataga cccgtcactg agtgggcagg tgagcataac cttgatagct actggcttca 240
aacggcagga tgaaccagaa ggccgcgtgt cgaagggtgg gcaacaaggt gagaatggcc 300
gacgcccatc cccagcanag ggcaacaaca cggtggaaat tccaaaattc ccgccaacaa 360
aaagggccct tccnnncttcc cacnattttg actggtcctg tctgcacctg tatga 415

<210> 16
<211> 744
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(744)
<223> n = A,T,C or G

<400> 16
aattcccggg tcgacccacg cgtcccgcgg acgcgtgggt ggaatatcac tggagggaac 60
gatctaacct tgacagaggt gaatgctgca gctgaagtaa tctatgatct tgttgaccct 120
ggtgcaaatc tgatttttgg ctctgttata gatccgtcat acactggtca agtgagcata 180
actctaattg caactggttt caaacgccag gaggaaagtg agagccggtc ttcacaggct 240
ggaggagaca agcaaccgcg gtcgctcggc tggttttctc ccacttccca ggaggaaggt 300
catgcattgc aaatcccana gttcctacag aggaaagggc gtccagggtt tcacgagtct 360
gaacacactt tggatcaatg ttttttcttgt catagtttgg tacgatgcag gtttggtttc 420
tgggtctctt aggtagcaag gtagaacaga tgttcctgaa cccgcacata ctaatctgtg 480
tgcaaacttc ngccgctgag taccattggc ttgggctgct ttgcttctca ngaacctgca 540
gtgaggtctc aatttgctag ttagtatgat taaaagtnaa gcgctgagac caaattatac 600
gttccgtgtg aatgattact tgctcnctgc cattttcttt tcaaaaaaaa aaaaaaaaaa 660
aaaaaggcgg cgctntanag gatccaagct tacttccccc gcatncgacn canagctntt 720
ntatagngtn acctaaattc aatc 744

<210> 17
<211> 230
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(230)
<223> n = A,T,C or G

<400> 17
ggctgctgag gaatcaagag aaaccatagc cactgccctg agggattcag atcttgtttt 60
cataacagct gggatgnnag ggggtgctgc tccaattgtt gcccagatat caaaggaagc 120
tggttatctt actgttggtg ttgtcaccta tccattcaat ttcgagggcc gtaagcgctc 180
tttacaggca agtatctgag ccccccttca ctcctgaatt agaattcaaa 230

<210> 18
<211> 318
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(318)
<223> n = A,T,C or G

<400> 18
caggcattgt gtggaatatc actgggggaa ctgacctaac tttgtttgag gtgaatgctg 60
cggccgaaat tatctacgac cttgtcgatc caaatgctaa tctgatattt ggtgccgtca 120
tagacccgtc actgagtggg caggtgagca taacctgata gctactggct tcaaacggca 180
ggatgaacca gaaggccgcg tgtcgaaggg tgggcaacaa agtgagaatg gccgacgccc 240
gtcccccgca gagggcagca gcacggtgga gttccagagt cctgcgacgt agaggancct 300
ctcgcttccc agagttga 318

<210> 19
<211> 471
<212> DNA
<213> Zea mays

<220>

<221> misc_feature
<222> (1)...(471)
<223> n = A,T,C or G

<400> 19
cgacgcccaa ggtgacgaat gctgtcagcc acgctgtgct acacggggga aacaatgcaa 60
anacattacc tgcctcactc ntgcttgctc ctgtaaatat aatgatngtc gctgctacat 120
natatttact cctgctgctg cttgaggcca ttattctgta cgtaaatgaa gccactacta 180
ctctcacaca gcatgcgccg gccgacgacg tacgtacgtg tattatatac gctctacccc 240
gtgagctttt gttcgagtga tacgtgatcc atccatgcat ggatgcttat gtatgtatat 300
gtgttagtcg tctcagggaa ccgggcanca naaggggtg ttgtattana tttacgtctt 360
ctggtgatta aataanaaag gggtatgttg gatgtgtgca aaaaaaaaaa aaaaanaaaa 420
aaaaaaaaaa aaaaaaaaag ggcggccgcc gactagtgag ctcgtcgacc c 471

<210> 20
<211> 1085
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(1085)
<223> n = A,T,C or G

<400> 20
cggctcgagg aaccctatta aaattggaga agttctgact cgtggattag gtacgggcgg 60
gaatccactt ttgggggaac aagctgcaga ggaatcaaga gatgctattg ctgatgctct 120
taaaggatca gatttggtgt ttataacggc tgggatgggt gggggaaccg ggtctggtgc 180
tgccccagtt gtagcccaaa tatcaaaaga ggcaggttac ttgactgtag gtgttgttac 240
ctatcccttc agttttgaag gacgtaagag atccttgcag gcctttgaag ccatcgaaag 300
gctgcagaaa aatgttgaca cmmttatagt gawtccmaat gmccgtctgc ttgacawagy 360
tratragcar atgcctcttc aaggatgctt tccgytttgc agatgacgtt ytmsggcaag 420
gagtmcaggg aatatcagac attatamctg tacctggact tkkcaaatgt ggattttgca 480
agatgtaaaa gctgtgatga aagactctgg gactgcaatg cttggagtag gtgtttccty 540
cggtaaaaaa ccgagcagaa gaagcagccg aacaggctac tttggctcct ttaattggat 600
cctctattca gtcaagctac tggggtagtg tataatatta ctggagggaa aggacataac 660
cctgcaggaa gtgracaggg tttytmaggt kgkgacyark ttggctgatc cttctgctaa 720
tattatattt ggggctgtcg ttgatgatcg ctacacgggg gagattcacg tgactatcat 780
tgcaactggc ttctcacagt cttttcagaa gaagttgcta acagatccaa gggcttgcaa 840
agctgcttga caaggtggct gagggccaag aaagcaaggc agtccctcct cccctcaagt 900
cctcaaacaa ggttgaatct agaccatccc cgcgaaagct ctttttttag ttgcatggtt 960
ctttttaccc tttttcattt ttccaattat tattattata ttatatnggc cgatcaaaaa 1020
aaaaaaaaaa ggcggccgcc gactagtgag ctcgtcgacc cgggaattaa ttccggaccg 1080
gtacc 1085

<210> 21
<211> 797
<212> DNA
<213> Glycine sp

<400> 21
ccagctggcg aaaggggatg tgctgcaagg cgattaagtt gggtacgcag ggttttccca 60
gtcacgacgt tgtaaaacga cggcagtgaa ttgaatttag gtgacactat agaagagcta 120
tgacgtcgca tgcacgcgta cgtaagctcg gaattcggct cgagaggcta ctttggctcc 180
tttaattgga tcctctattc agtcagctac tggggtagtg tataatatta ctggaggaaa 240

ggacataacc ctgcaggaag tgaacagggt ttctcaggtt gtgactagtt tggctgatcc 300
ttctgctaat attatatttg gggctgtcgt tgatgatcgc tacactgggg agattcacgt 360
gactatcatt gcaactggct tctcacagtc ttttcagaag aagttgctaa cagatccaag 420
ggctgcaaag ctgcttgaca aggtggctga gggccaagaa agcaaggtag tccctcctcc 480
cctcaagtcc tcaaacaagg ttgaatctag accatccccg cgaaagcttt ttttttagt 540
tgcatggttc tttttaccct ttttcatttt tccaattatt attattatat tatattggcc 600
gatcaaaaaa aaaattatta tattatattg taggacacaa tgatcttgat gcttaattaa 660
gtgagatatc attctcttga tgttaaaaaa aaaaaaaaag ggcggccgcc gactagtgag 720
ctcgtcgacc cgggaattaa ttccggaccg gtacctgcag gcgtaccagc tttccctata 780
gtgagtcgta ttagagg 797

<210> 22
<211> 714
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(714)
<223> n = A,T,C or G

<400> 22
aattcggctc gagacggctg cgagaagacg acagaagggg gttaccgtta tcatgcaagc 60
tgataatggg gcctctgaag ttcttgttcc gttattataa aactgagtcc ttcactctct 120
ctcgaaccag ctcacagaaa caatgatctc ctacgccgac atgctcaagg gatcacatgg 180
atgtcaacaa cttcaactat cctccattgt cagagatgta aactacagct gtggctcgtg 240
tggttatgag ctgaacttga actccagcaa ccgcaacact tgttctctca ttgactcaaa 300
gtccataaag agaggcatca tctccttctt ctccgtggat gagagcaggt tcactcagat 360
ccagcaactt cactggcctt cttggatgcc ctttttcaac tccaagcgcc aaagaaccaa 420
gctttttttgc cgcagctgtg ggaaccacct tggctatgct tacactttgc ctctcaatct 480
caatcccggg atggcatctc tgatgattca gaatctatga tatcaaacta accgctttgt 540
taccttcttt ctgcgaggaa ccaagtcaaa agttaganga tatgggcaag gtttgagact 600
gcatcttcct ccactcttgg tggtctaatt cttgaaaggg acagaaacat attcatcagt 660
tcttggttgg ttggaatgng aattaatgna ttctaccttt tgacattatg aagg 714

<210> 23
<211> 525
<212> DNA
<213> Glycine sp

<400> 23
cgggctcgag attactggag gaaaggacat aaccctgcag gaagtgaaca gggtttctca 60
ggttgtgact agtttggctg atccttctgc taatattata tttggggctg tcgttgatga 120
tcgctacact ggggagattc acgtgactat cattgcaact ggcttctcac agtcttttca 180
gaagaagttg ctaacagatc caagggctgc aaagctgctt gacaaggtgg ctgagggcca 240
agaaagcaag gtagtccctc ctcccctcaa gtcctcaaac aaggttgaat ctagaccatc 300
cccgcgaaag ctctttttt agttgcatgg ttctttttac ccttttcat ttttccaatt 360
attattatta tattatattg gccgatcaaa aaaaaaatta ttatattata ttgtaggaca 420
caatgatctt gatgcttaat taagtgagat atcattctct tgatgttctt tcccctccaa 480
aaaaaaaaaa aaagggcggc cgccgactag tgagctcgtc gaccc 525

<210> 24
<211> 1083
<212> DNA
<213> Glycine sp

<400> 24
cggctcgagg aaccctatta aaattggaga agttctgact cgtggattag gtacgggcgg 60
gaatccactt ttgggggaac aagctgcaga ggaatcaaga gatgctattg ctgatgctct 120
taaaggatca gatttggtgt ttataacggc tgggatgggt gggggaaccg ggtctggtgc 180
tgccccagtt gtagcccaaa tatcaaaaga ggcaggttac ttgactgtag gtgttgttac 240
ctatcccttc agttttgaag gacgtaagag atccttgcag gcctttgaag ccatcgaaag 300
gctgcagaaa aatgttgaca cacttatagt gattccaaat gaccgtctgc ttgacatagc 360
tgatgagcag atgcctcttc aggatgcttt tccgtcttgc agatgacgtt ctacggcaag 420
gagtacaggg aatatcagac attatamctg wcctggactt gtcaatgtgg atttttgcag 480
atgtaaaagc tgtgatgaaa gactctggga ctgcaatgct tggagtaggt gtttcctccg 540
gtaaaaaccg agcagaagaa gcagccsaac aggctacttt ggctycttta attggatcct 600
ctatttcagt cagctactgg gggtagtgta taatattact ggaggaaagg acataaccct 660
scaggaagtg aacagggktt ctcaggttgt gactaagttt ggctgatcct tctgctaata 720
ttatatttgg ggctgtcgtt gatgatcgct acacgggggа gattcacgtg actatcattg 780
caactggctt ctcacagtct tttcagaaga agttgctaac agatccaagg gctgcaaagc 840
tgcttgacaa ggtggctgag ggccaagaaa gcaaggcagt ccctcctccc ctcaagtcct 900
caaacaaggt tgaatctaga ccatccccgc gaaagctctt ttttttagttg catggttctt 960
tttacccttt ttcatttttc caattattat tattatatta tattggccga tcaaaaaaaa 1020
aaaaaaaggg cggccgccga ctagtgagct cgtcgacccg ggaattaatt ccggaccggt 1080
acc 1083

<210> 25
<211> 1335
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(1335)
<223> n = A,T,C or G

<400> 25
cggctcgagg cccagaacaa caaaaattgc tcctcaacgc ctaagtcgtc gtttcggttc 60
ggtgagatgc tcctacgctt acgtagataa cgccaaaatt aaggttgtcg gcatcggcgg 120
tggcggcaac aatgccgtta atcgcatgat cggaagtggt ttgcagggtg tagacttcta 180
tgcgataaat accgatgctc aggcactatt aaattctgct gctgagaacc ctattaaaat 240
tggagaagtt ctgactcgtg gattaggtac aggtgggaat ccacttttgg gggaacaagc 300
tgcggaggaa tccagagatg ctattgctga tgctcttaaa ggatcagatt tggtatttat 360
aacggctggg atgggtgggg gaaccgggtc ttggtgctgc cccagttgta gnccaaatat 420
caaaagaggc aggntacttt gactgtaggt gttggtacct atcccttcag ttttgaagga 480
cgtaagagat gcttgcaggc ctttgaagcc atcgaaaggc tgcagaaaaa tgttgcacac 540
ttatagttat tccaaatgat cgtctgcttg acatancttg atgaaccaga tgcctattca 600
aggatgcttt ycgytytkca rawkatgtty tamcgsaarg sgkacaggga atatcaagac 660
attwtaacag gtacctggac ttgtmaatgt agattttgct gatgtaaaam ctgkgataaa 720
gacttctggg actgcaatgc ttggtgtagg tgtttcatcc ggtaaaaccg accagaagaa 780
gcagcagaac agggctactt tggctccttt aattggatca tctattcagt cagctactgg 840
ggtagtgtat aatattactg gaggaaagga cataaccctg caggaagtga acagggtttc 900
tcaggtggtg actagtttgg ctgatccttc tgctaatatt atatttggag cttgttgttg 960
atgatcgctt acactgggga gattcacgtg actataattg caactggctt ctcacagtct 1020
tttcagaaga agttgctaac agatccaagg gctgcaaagc tgcttgacaa agtggctgag 1080
ggccaagaaa gcaaggcagt ccctcctccc cccaagtcct caatcaaggt tgaatctaga 1140
ccatccccgc gaaagctctt tttgtagttg catggttctt ttacctttt ctttttttcca 1200
attattatat tgtaagtcat tctgtagtac aatgatcttg atgcttaatt tagtgagata 1260
tcattctctt gatgttaaaa aaaaaaaaaa aaaaaaaaaa aaagggcggc cgccgactag 1320

tgagctcgtc gaccc 1335

<210> 26
<211> 902
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(902)
<223> n = A,T,C or G

<400> 26
aattcggctc gagtaccagg gttggtgaat gtagattttg ctgatgttcg ggctataatg 60
gccaatgcag gttcttcact aatggggata ggaactgcaa ctggaaaatc aagggcaaga 120
gatgctgcat taaatgccat ccagtcacct ttactggata ttggtatara gagggctact 180
kgaattgttt ggaacawaac tggtgggact gatctgrcct tgtttgaggt aaacacggca 240
gcagaggtta tttatgacct cgtggaccct actgctaatt taatatttgg agcagtaata 300
gatccatcac tcagtggtca agtgagcata acattaattg cttactgrat tcaaagcgyc 360
aagaggagag tgaagggagg cctctgcagg ccagtcaact cactcaagca gacacaacct 420
tcggcaccaa ttggcggtct tcctctttca ctgatggtgg tttgtttgag ataccagaat 480
tcctaaagaa garaggaggt tcacgctatc cgagggcgta atctttttca tcctaatttc 540
ttttgatccc ttgcatttct tcaccttgg atatacatag caattggtct agttcttarg 600
tccctgtctt gsccttttc ggatttwrkc aaragttgkg katacagttk gttcatgaaa 660
gtttattact tyccactgkc cagacttatg ggkctaaacc gganggtatt ksarcatgga 720
tgcttttctt ggcatatttg aattagttta ttagcttgta cagagatttc agtaatgctg 780
agagcttgtt atagttcttt ggcatgttat agaaaattca ttattattaa aaaaaaaaaa 840
aaaggcggcc gccgactagt gagctcgtcg acccgggaat tnattccgga ccggtacctg 900
ca 902

<210> 27
<211> 856
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(856)
<223> n = A,T,C or G

<400> 27
aattcggctc gagattggtg aaccgtagac tttgctgatg ttcgagctat aatggccaat 60
gcagggtctt cacttatggg gataggaact gcaactggga aaacaagggc marggawgct 120
gcattaaatg ctatccagtc mccctttact ggatatttgg tataraaagg gctactggaa 180
ttgtatggaa cataacyggk ggaagtgatt tgaccttgtt tgaaggtaaa tgttgcasca 240
raagttatat atgmccttgt ggmccccact gstaatttaa tatttgggsc agwaatagat 300
ccatcactcc agtgggcaag taagcatamm wtaatcgcaa ctggattcaa gcgtcaagag 360
gaaaagtgaa gggagaccct atgcaggcca gtcaactcac acaaggagat nccgttggta 420
tcaatcggcg atyttctact ttcactgatg gtagcttttg ttggagatcc ctggaattct 480
taaagaagaa ggggcgctca cgttatccaa gagtttaata ctcttttccc caactcctta 540
atccctcctt gcatctcttt mccaascaat ttttagggat acaaatctca tcagtctaag 600
gtattagatc acggtttttg ccccttttt cattttagg ttcgcattgt gcantamagt 660
tgttcattga aagcgaagtt actttccaaa accgttgttt tctgarttga aggcttggtt 720
ggcatgtttt wataagttta ttagcttgta tttttgtnca gagaataata tatcagtaat 780
ggtcagtgct tgttataaan ccncnaaaaa aaaaaaaaaa aaaagggcgg ccgccgacta 840

gtgagctcgt cgaccc 856

<210> 28
<211> 1060
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(1060)
<223> n = A,T,C or G

<400> 28
aattcggctc gaggtcacaa ccccttttc atttgaaggg cgaagaaggg cagttcaagc 60
acaagaagga attgctgcat taagagataa tgttgataca ctgatagtta ttccaaatga 120
caaactgctg actgcagttt ctcaatctac ccctgtaact gaagcattca atctggctga 180
tgatattctt agacaaggtg ttcgtggtat atctgatatt attacgatac caggattggt 240
gaatgtagac tttgcagatg ttcgagctat aatggccaat gcaggttctt cacttatggg 300
gataggaact gcaactggca aaacaagggc aagagacgct gcattaaatg ctatccagtc 360
acctttacta gatattggta tagaaagggc taccggaatt gtatggaaca taactggtgg 420
aaagwgattt gaccttgttt gaggtaaatg ctgcagcaga agttatatat gaccttgtgg 480
accccactgy taatttaata tttggagcag taatagatcc atcactcagt ggtcaagtaa 540
gcatcacatt aattgcaact ggattcaagc cgtcaagagg aaaagtgaag ggagacctat 600
acaggccarc caatttacac aaggagatac ggttggtatc aaccsgcgat cttcctcttt 660
cactgatggk agctttgttg agaycccтga attcttaaag aaraaggggc gctcaygtta 720
tccgagagct taatactctt ctccccaatt tcttaatccc ttgatttctt tacaaagtaa 780
tttttaggga tacaaatctc atcagtctag gtattagatc ccgttttgcc ccttttttt 840
ttcattttta ggttcgcatt gggcatactg ttgttcaaga aagcaaagta ctttcaaaac 900
cgttgtttac tgagtcgagg cttgttggca ggttttaata agtttattag cttgtatttt 960
ttgtacagag aatatatcag taatggtcag ggcttgttat nnnannnccn nnnannaaan 1020
aaaaaaaaag gcggccgccg actagtgagc tcgtcgaccc 1060

<210> 29
<211> 727
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(727)
<223> n = A,T,C or G

<400> 29
atctcncaaa atgcatgncn ctgtgtgtgg catatattca aaatgacttg gcccagggtg 60
gggttttant ttgctcttag aaaatgtgtt gagcctgcac atanaagatt ggagttgttg 120
attctcagtg gattgttcac caaggtattc cctcactagg gaatcagggt gantctcaaa 180
caggaaagcn ccatggcagg ggntgaggga ncggtgtana aaggagtggc catgttccag 240
agtcggtggc aaatgctgaa tacgcgtatc acaactccat tggaattgat acatctaatt 300
ccactgctca ttaggtgact tcggcctaag ttgacttgta aacatattgt tactaccctt 360
agccttacgc gtagaatttt cccttaaaaa aaaaaatata ttcctatgta acgttacgta 420
catgcaatgc aatcacaata tagagtccta gctagggacc aaacatcatt tcgatgtaga 480
aattgctgta cttaacagtg agtaaatcta gtgaagagaa ttattattgc tgctaacgaa 540
ggtgcttata ggaaatggaa atgctagtga atccttaaat tggaggctga caacgaagtt 600
ctttagggtt tttgggatta aagaaaacga aatgtcataa ttatcatacc cttgggatga 660
ggagacagga ctattactat aaaaaaaaaa aaaaagggcg gccgccgact agtgagctcg 720

tcgaccc 727

<210> 30
<211> 1185
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(1185)
<223> n = A,T,C or G

<400> 30
cggctcgagc tggaatgggt gggggaactg gcacaggtgg agctccaatt attgctagta 60
ttgcaaagtc aatgggtata ttgacggttg gtattgtcac cacccctttc tcgtttgaag 120
ggagaaagag atctattcaa gcccaagaag gaattacagc cttaagagat aatgttgaca 180
cgcttatagt tattccaaat gacaagctac taacggcagt ttctcaatct acccctgtaa 240
ctgaagcatt caatctggct gatgatattc ttcgacaggg tgttcgtggc atatctgata 300
ttattacaat accagggttg gtgaatgtag attttgctga tgttcgggct ataatggcca 360
atgcaggttc ttcactaatg gggataggaa ctgcaactgg aaaatcaagg gcaagagatg 420
ctgcattaaa tgccatccag tcaccwttmc tggatattgg tatagararg gctactggaa 480
ttgtttggaa cawaactggk gggactgatc ttgaccttgt ttgaggtaaa cacggcarca 540
rraggttatt tatgacctcg tggaccctac tgctaattta atatttggag cagtaataga 600
tccatcactc agtggkcaag tgagcataac attaattgct actggattca agcgtcaaga 660
ggarartgaa rggaggcctn tgcaggccag tcaactcact caagcagaca caaccttcgg 720
caccaattgg cggtcttcct ctttcactga tggtggtttg tttgagatac cagaattcct 780
aaagaagaga ggaggttcac gctatccgag ggcgtaatct ttttcatcct aatttctttg 840
atcccttgca tttcttcacc cttggatata catagcattg gtctagttct taggtccctg 900
tcttgccctt tttcggattt tagtcagagt tgtgtataca gtttgttcat gaaagtttat 960
tacttcccac tgtccagact tatgggtcta accggaggta ttgcagcatg gatgcttttc 1020
ttggcatatt tgaattagtt tattagcttg tacagagatt tcagtaatgc tgagagcttg 1080
ttatagttct ttggcatgtt atagaaaatt cattattatt attcatcccn ccaaaaaaaa 1140
aaaaaaaaaa aaagggcggc cgccgactag tgagctcgtc gaccc 1185

<210> 31
<211> 700
<212> DNA
<213> Glycine sp

<220>
<221> misc_feature
<222> (1)...(700)
<223> n = A,T,C or G

<400> 31
aattcggctc gagattgtca ccaccccttt ctcgtttgaa gggagaaaga gatctattca 60
agcccaagaa ggaattacag ccttaagaga taatgttgac acgcttatag ttattccaaa 120
tgacaagcta ctaacggcag tttctcaatc tacccctgta actgaagcat tcaatctggc 180
tgatgatatt cttcgacagg gtggtccgtg gcatatctga tattattaca ataccagggt 240
tggtgaatgt agattttgct gatgttcggg ctataatggc caatgcaggt tcttcactaa 300
tggggatagg aactgcaact ggaaaatcaa gggcaagaga tgctgcatta aatgccatcc 360
agtcaccttt actggatatt ggtatagaga gggctactgg aattgtttgg aacataactg 420
gtgggactga tctgccttgt ttgaggtaaa cacngcagca ganggtattt atgacctcgn 480
ggccctactg ctaattaata tttggagcag aatagatcca tcctcatggc aagtgacata 540
cattnantgc tctggattca agcgtcaaga ngagaagtga agggangcct ttgcaggcca 600

gcactcactc agcagacaca accttngnac caattggcgg cttcctcttt cactgatggg 660
nggttggttg agatncnana attcctaaag aaaaanagag 700

<210> 32
<211> 1425
<212> DNA
<213> Arabidopsis sp

<400> 32
ccacgcgtcc ggaggaagta aacaatggcg ataattccgt tagcacagct taatgagcta 60
acgatttctt catcttcttc ttcgtttctt accaaatcga tatcttctca ttcgttgcat 120
agtagctgca tttgcgcaag ttctagaatc agtcaattcc gtggcggctt ctctaaacga 180
agaagcgatt caacaaggtc taagtcgatg cgattgaggt gttccttctc tccgatggaa 240
tctgcgagaa ttaaggtgat tggtgtcggt ggtggtggta acaatgccgt taaccggatg 300
atttcaagcg gtttacagag tgttgatttc tatgcgataa acacggattc gcaagctctg 360
ttacagtttt ctgctgagaa cccacttcaa attggagaac ttttaactcg tgggcttggc 420
actggtggaa acccgcttct tggagaacaa gctgcagaag aatcaaaaga tgcaattgct 480
aatgctctta aaggatcaga ccttgttttc ataactgctg gtatgggtgg tggaacaggg 540
tctggtgctg cacctgtggt agctcagatt tcgaaggatg ctggttattt gactgttggt 600
gttgttacct atccgtttag ctttgaagga cgtaaaagat ctttgcaggc actggaagct 660
attgaaaagc tccaaaagaa tgttgatacc cttatcgtga ttccaaatga tcgtctgcta 720
gatattgctg atgaacagac gccacttcag gacgcgtttc ttcttgcaga tgatgtttta 780
cgccaaggag tacaaggaat ctcagatatt attactatac ctggactagt caatgtggat 840
tttgcagatg tgaaggcagt catgaaagat tctggaactg caatgctcgg ggtaggtgtt 900
tcttccagca aaaaccgggc agaagaagca gctgaacaag caactttggc tccattgatc 960
ggatcatcca tacaatcagc tactggtgtc gtctacaaca tcactggtgg aaaagacata 1020
actttgcagg aagtgaaccg agtatcacag gtcgtgacaa gtttggcaga cccatcggcc 1080
aacatcatat ttggagctgt tgtggatgat cgctacaccg gagagattca tgtaacgata 1140
atcgccacag gcttctctca gtcattccag aagacacttc tgactgatcc aagagcagct 1200
aaactccttg acaaaatggg atcatcaggt caacaagaga acaaaggaat gtctctgcct 1260
caccagaagc agtctccatc aactatctct accaaatcgt cttctccccg tagacttttc 1320
ttctagtttt ctttttttcc ttttcggttt caagcatcaa aaatgtaacg atcttcaggc 1380
tcaaatatca attacatttg attttcctcc aaaaaaaaaa aaaaa 1425

<210> 33
<211> 1611
<212> DNA
<213> Arabidopsis sp

<400> 33
tgttgttgcc gctcagaaat ctgaatcttc tccaatcaga aactctccac ggcattacca 60
aagccaagct caagatcctt tcttgaacct tcacccggaa atatctatgc ttagaggtga 120
agggactagt acaatagtca atccaagaaa ggaaacgtct tctggacctg ttgtcgagga 180
ttttgaagag ccatctgctc cgagtaacta caatgaggcg aggattaagg ttattggtgt 240
gggaggtggt ggatcaaatg ctgtgaatcg tatgatagag agtgaaatgt caggtgtgga 300
gttctggatt gtcaacactg atatccaggc tatgagaatg tctcctgttt tgcctgataa 360
taggttacaa attggtaagg agttgactag gggtttaggt gctggaggaa atccagaaat 420
cggtatgaat gctgctagag agagcaaaga agttattgaa gaagctcttt atggctcaga 480
tatggtcttt gtcacagctg gaatgggcgg tggaactggc actggtgcag cccctgtaat 540
tgcaggaatt gccaaggcga tgggtatatt gacagttggt attgccacaa cgcctttctc 600
gtttgagggt cgaagaagaa ctgttcaggc tcaagaaggg cttgcatctc tcagagacaa 660
tgttgacact ctcatcgtca ttccaaatga caagttgctt acagctgtct ctcagtctac 720
tccggtaaca gaagcattta atctagctga tgatatactc cgtcaggggg ttcgtgggat 780
atctgatatc attacgattc ctggtttggt gaatgtggat tttgctgatg tgagagctat 840
aatggcaaat gcggggtctt cattgatggg aataggaact gcgacaggaa agagtcgggc 900

aagagatgct gcgctaaatg caatccaatc ccctttgtta gatattggga ttgagagagc 960
cactggaatt gtttggaaca ttactggcgg aagtgacttg acattgtttg aggtaaatgc 1020
tgctgcggaa gtaatatatg atcttgtcga tccaactgcc aatcttatat tcgtgctgtt 1080
gtagatccag ccctcagcgg tcaagtaagc ataaccctga tagctacggg tttcaaacga 1140
caagaagagg gagaaggacg aacagttcag atggtacaag cagatgctgc gtcagttgga 1200
gctacaagaa gaccctcttc ttcctttaga gaaagcggtt cagtggagat cccagagttc 1260
ttgaagaaga aaggcagctc tcgttatccc cgagtctaaa gcccaatcta atcactaccc 1320
tgcacactgc agcaataaca aacgtgtgtg tactggtagt ctggtactgc cttctgggat 1380
acagcaagat gtgttgatgt atgatcaaga atctgtgtgg gtgtgtatat gttctgtcac 1440
tgcctctggt cgtgttcttg aataggttgt tttagaaatc ggagtttctc tctatgtcac 1500
ttccaaaaca aaaaaggaga agaagaatca cacttctcga accataaaca tacttataag 1560
attatgagag ttttagcaga aattattgtc aaaaaaaaaa aaaaaaaaaa a 1611

<210> 34
<211> 299
<212> DNA
<213> Arabidopsis sp

<220>
<221> misc_feature
<222> (1)...(299)
<223> n = A,T,C or G

<400> 34
agtaattgaa aaatgacaca tacctctttt nctacaatca aagtttataa ctaagaagaa 60
caaatgcaat agtaataata aagatttgag gacatactgt gacaaagacc atatctgagc 120
cataaagagc ttctncaata actgctttgc tctctctagc agcattcata ccgatttcng 180
gatttcctcc agcacctaaa cccctagtca actccttacc aatttgtaac ctattatcag 240
gcaaaacagg agacattctc atagcctgga tatcagtgtt gacaatccag aactccaca 299

<210> 35
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 35 25
cttgccgcaa aacatcatcc gcgag

<210> 36
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 36 28
cacccacccc gagcatcgca gttccgga

<210> 37
<211> 41
<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 37
ttggtgtcgg tggtggtggt aacaatgccg ttaaccggat g 41

<210> 38
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 38
ggtaacaatg ccgttaaccg gatgatttca agcggtttac a 41

<210> 39
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 39
gctgctcttg gatcagtcag aagtgtcttc tggaatgact g 41

<210> 40
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 40
tttgtcaagg agtttagctg ctcttggatc agtcagaag 39

<210> 41
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 41
cagcaccaga tcctgtacct ccacccat 28

<210> 42
<211> 30
<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 42 30
ctgaatgggt atgtgacaac accaacagtc

<210> 43
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 43 28
gctgatgatg tattacgcca aggtgtcc

<210> 44
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 44 28
ctggaactgc tatgcttgga gttggggt

<210> 45
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 45 34
tcgaaactga atggataggt gacaacacca acag

<210> 46
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 46 36
tggatcgaca aggtcgtaga taatttcggc cgcagc

<210> 47
<211> 56
<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 47
cgcgatttaa atggcgcgcc ctgcaggcgg ccgcctgcag ggcgcgccat ttaaat 56

<210> 48
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 48
tcgaggatcc gcggccgcaa gcttcctgca gg 32

<210> 49
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 49
tcgacctgca ggaagcttgc ggccgcggat cc 32

<210> 50
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 50
tcgacctgca ggaagcttgc ggccgcggat cc 32

<210> 51
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 51
tcgaggatcc gcggccgcaa gcttcctgca gg 32

<210> 52
<211> 36
<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 52
tcgaggatcc gcggccgcaa gcttcctgca ggagct 36

<210> 53
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 53
cctgcaggaa gcttgcggcc gcggatcc 28

<210> 54
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 54
tcgacctgca ggaagcttgc ggccgcggat ccagct 36

<210> 55
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 55
ggatccgcgg ccgcaagctt cctgcagg 28

<210> 56
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> c-myc Tag

<400> 56
Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> 57
<211> 37

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 57
acgtgcggcc gcatggcgat aattccgtta gcacagc 37

<210> 58
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 58
actgataagc ttttgctcga agaaaagtct acgggga 37

<210> 59
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 59
acgtgcggcc gcatggcgat aattccgtta gcacagc 37

<210> 60
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 60
cgtcctgcag gctacaagtc ttcctcactg ataagctttt gctc 44

<210> 61
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 61
acgtcctgca ggatggcgat aattccgtta gcacagc 37

<210> 62
<211> 38

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 62
acgtgcggcc gcctagaaga aaagtctacg gggagaag 38

<210> 63
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 63
acgtaagctt tccttctctc cgatggagtc tg 32

<210> 64
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 64
acgtctgcag cttttcaatg gcttcaagtg cct 33

<210> 65
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 65
acgtctgcag aagaacgtgg ataccctcat cgtg 34

<210> 66
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 66
acgtgcggcc gcctagaaga acaatctacg gggagaag 38

<210> 67
<211> 244

<212> DNA
<213> Artificial Sequence

<220>
<223> Prrn/NEP/Gene10leader/14 amino acid from GFP
fusion

<400> 67
gaattcggta cccccgtcgt tcaatgagaa tggataagag gctcgtggga ttgacgtgag 60
ggggcaggga tggctatatt tctgggagcg aactccgggc gaatactgaa gcgcttggat 120
acaagttatc cttggaagga aagacaattc cggatcctct agaaataatt ttgtttaact 180
ttaagaagga gatataccca tgggtaaagg agaagaactt ttcactggag ttgtcccaag 240
catg 244

<210> 68
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 68
gtcctgcagg atggccacca tctcaaaccc ag 32

<210> 69
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Oligonucleotide

<400> 69
tagcggccgc ctatataaag gagctaaaag aacagcc 37

<210> 70
<211> 669
<212> DNA
<213> Arabidopsis sp

<400> 70
atggcgatta gtccgttggc acagcttaac gagctaccag tctcttcctc gtttsttgcg 60
acatcccact cgctgcacag taccagaatc agtggcgggc ttctcaaaac aaaggtttaa 120
gcaaacacgg ttgagatgct ccttctctcc gatggagtct gcgaggatta aggtggttgg 180
tgtcggcggt ggtggtaaca atgccgtcaa tcgcatgatt tccagcggct tacagagtgt 240
tgatttctat gcgataaaca cggactctca agctctcttg cagtcttctg cgcagaaccc 300
tcttcaaatt ggagagctcc taactcgtgg ccttgggact ggtgggaacc cgcttctagg 360
agaacaagct gctgaggaat ctaaagacgc gattgctaat gctcttaaag gatctgacct 420
tgytttcatt actgctggta tgggtggtgg cactggctcc ggtgctgctc ctgttgttgc 480
tcagatctcc aaagacgctg gttatttgac cgttggtgtt gttacctatc ccttcagctt 540
cgaaggtcgt aaaagatctt tgcaggcact tgaagccatt gaaaagctgc agaakaacgt 600
ggataccctc atcgtgatac caaatgatcg tctcctagat attgctgatg aacagacgcc 660
tcttcaaga 669

<210> 71
<211> 646
<212> DNA
<213> Nicotiana sp

<400> 71
ggccctctag atgcatgctc gagcggccgc cagtgtgatg gatatctgca gaattcgccc 60
ttaagcagtg gtaacaacgc agagtacgcg ggggtaaacc aaacagacag agagagcaga 120
aacagcaatg gccaccatct caaacccagc agagatagca gcttcttctc cttcctttgc 180
tttttaccac tcttccttta ttcctaaaca atgctgcttc accaaagctc gccggaaaag 240
cttatgtaaa cctcaacgtt tcagcatttc aagttcattt actccttttg attctgctaa 300
gattaaggtt atcggcgtcg gtggcggtgg taacaatgcc gttaaccgta tgattggcag 360
tggcttacag ggtgttgact tctatgctat aaacacggat gctcaagcac tgctgcagtc 420
tgctgctgaa aatccacttc aaattggaga gcttctgact cgtgggcttg gtactggtgg 480
caatcctctt ttaggggaac aggcagcaga ggagtcgaag gaagccattg caaattctct 540
aaaaggttca gatatggtgt tcataacagc aggaatgggt ggaggtacag gatctggtgc 600
tgaagggcga attccagcac actggcggcc gttactagtg gatccg 646

<210> 72
<211> 325
<212> PRT
<213> Glycine sp

<400> 72
Gly Ser Arg Pro Arg Thr Thr Lys Ile Ala Pro Gln Arg Leu Ser Arg
1 5 10 15
Arg Phe Gly Ser Val Arg Cys Ser Tyr Ala Tyr Val Asp Asn Ala Lys
20 25 30
Ile Lys Val Val Gly Ile Gly Gly Gly Gly Asn Asn Ala Val Asn Arg
35 40 45
Met Ile Gly Ser Gly Leu Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr
50 55 60
Asp Ala Gln Ala Leu Leu Asn Ser Ala Ala Glu Asn Pro Ile Lys Ile
65 70 75 80
Gly Glu Val Leu Thr Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu
85 90 95
Gly Glu Gln Ala Ala Glu Glu Ser Arg Asp Ala Ile Ala Asp Ala Leu
100 105 110
Lys Gly Ser Asp Leu Val Phe Ile Thr Ala Gly Met Gly Gly Gly Thr
115 120 125
Gly Ser Gly Ala Ala Pro Val Val Ala Gln Ile Ser Lys Glu Ala Gly
130 135 140
Tyr Leu Thr Val Gly Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg
145 150 155 160
Lys Arg Ser Leu Gln Ala Phe Glu Ala Ile Glu Arg Leu Gln Lys Asn
165 170 175
Val Asp Thr Leu Ile Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala
180 185 190
Asp Glu Gln Met Pro Leu Gln Asp Ala Phe Pro Phe Ala Asp Asp Val
195 200 205
Leu Arg Gln Gly Val Gln Gly Ile Ser Asp Ile Ile Thr Val Pro Gly
210 215 220
Leu Val Asn Val Asp Phe Ala Asp Val Lys Ala Val Met Lys Asp Ser
225 230 235 240
Gly Thr Ala Met Leu Gly Val Gly Val Ser Ser Gly Lys Asn Arg Ala

```
                245                 250                 255
Glu Glu Ala Ala Glu Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser
            260                 265                 270
Ile Gln Ser Ala Thr Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp
        275                 280                 285
Ile Thr Leu Gln Glu Val Asn Arg Val Ser Gln Val Val Thr Ser Leu
    290                 295                 300
Ala Asp Pro Ser Ala Asn Ile Ile Phe Gly Ala Val Val Asp Asp Arg
305                 310                 315                 320
Tyr Thr Gly Glu Ile
                325
```

<210> 73
<211> 357
<212> PRT
<213> Zea mays

<220>
<221> VARIANT
<222> (1)...(357)
<223> Xaa = Any Amino Acid

```
<400> 73
Asp Leu Val Phe Ile Thr Ala Gly Met Gly Gly Gly Thr Gly Ser Gly
 1               5                  10                  15
Ala Ala Pro Val Val Ala Gln Ile Ser Lys Glu Ala Gly Tyr Leu Thr
            20                  25                  30
Val Gly Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser
        35                  40                  45
Val Gln Ala Leu Glu Ala Leu Glu Lys Leu Glu Lys Ser Val Asp Thr
    50                  55                  60
Leu Ile Val Ile Pro Asn Asp Lys Leu Leu Asp Val Ala Asp Glu Asn
65                  70                  75                  80
Met Pro Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln
                85                  90                  95
Gly Val Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn
            100                 105                 110
Val Asp Phe Ala Asp Val Lys Ala Val Met Lys Asn Ser Gly Thr Ala
        115                 120                 125
Met Leu Gly Val Gly Val Ser Ser Ser Lys Asn Arg Ala Gln Glu Ala
    130                 135                 140
Ala Glu Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Glu Ala
145                 150                 155                 160
Ala Thr Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu
                165                 170                 175
Gln Glu Val Asn Lys Val Ser Gln Ile Val Thr Ser Leu Ala Asp Pro
            180                 185                 190
Ser Ala Asn Ile Ile Phe Gly Ala Val Val Asp Asp Arg Tyr Thr Gly
        195                 200                 205
Glu Ile His Val Thr Ile Ile Ala Thr Gly Phe Pro Gln Ser Phe Gln
    210                 215                 220
Lys Ser Leu Leu Ala Asp Pro Lys Gly Ala Arg Ile Val Glu Ser Lys
225                 230                 235                 240
Glu Lys Ala Ala Thr Leu Ala His Lys Ala Ala Ala Ala Ala Val Gln
                245                 250                 255
```

```
Pro Val Pro Ala Ser Ala Trp Ser Arg Arg Leu Phe Ser Xaa Glu Ala
                260                 265                 270
His Leu Val Asn Arg Asp Ser Xaa Cys Ile Arg Phe Ala Phe Ser Val
        275                 280                 285
Leu Arg Ala Val Pro Lys Val Ile Phe Gly Tyr Leu Glu Ile Tyr Ser
    290                 295                 300
Leu Gly Xaa Cys Ser Val Val Val Glu Xaa Val Ser Val Tyr Val Ser
305                 310                 315                 320
Leu Leu Cys Phe Met Phe Leu Arg Ile Xaa Arg Xaa Gly Xaa Glu Lys
                325                 330                 335
Cys Ser Ala Thr Gln His Xaa Thr Val Xaa Lys Ile Phe Asp Cys Phe
            340                 345                 350
Ile Ala Ala Thr Cys
        355
```

INTERNATIONAL SEARCH REPORT

Inte .tional Application No
PCT/US 99/28103

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/82 C12N15/29 C07K14/415 C12N5/10 A01H1/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	WO 98 00436 A (UNIV NEVADA) 8 January 1998 (1998-01-08) abstract; claims 1,8 page 2, line 30 -page 3, line 24 page 6, line 3 - line 16 page 6, line 33 -page 7, line 4 page 7, line 21 -page 8, line 31 page 10, line 26 - line 36 --- -/--	16,17, 24-26,30 1-11, 18-23

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 April 2000	17/04/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Ceder, O

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int. ..tional Application No
PCT/US 99/28103

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	STREPP ET AL: "Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 95, 1 April 1998 (1998-04-01), pages 4368-4373, XP002083808 ISSN: 0027-8424 cited in the application	16, 24-26,30
A	abstract page 4368, left-hand column	6-12
X	SVAB ET AL: "High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 90, February 1993 (1993-02), pages 913-917, XP002106110 ISSN: 0027-8424 cited in the application	27
A	abstract	1
A	YU ET AL.: "Inhibition of assembly of bacterial cell division protein FtsZ by the hydrophobic dye 5,5'-bis-(8-anilino-1-naphthalenesulfonate)" THE JOURNAL OF BIOCHEMICAL CHEMISTRY, vol. 273, no. 17, 24 April 1998 (1998-04-24), pages 10216-10222, XP002134449 cited in the application abstract page 10221, right-hand column, line 5 - line 7 page 10221, right-hand column, line 25 -page 10222, left-hand column, line 3	6,12-14
A	WO 93 10253 A (ESCAGENETICS INC) 27 May 1993 (1993-05-27) abstract; claim 1 page 5, line 14 - line 17	1,2

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte ..ional Application No
PCT/US 99/28103

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9800436 A	08-01-1998	AU 3644497 A	21-01-1998
		CA 2259209 A	08-01-1998
		EP 0912596 A	06-05-1999
		US 5981836 A	09-11-1999
WO 9310253 A	27-05-1993	US 5344775 A	06-09-1994
		AU 3139593 A	15-06-1993